REMARKS

Specification

The specification is objected to because Figures 14 and 15 each comprise a number of sequences but there is no indication in the figure or the description of the figure as to the associated SEQ ID NOs.

Applicant submits that the specification had been amended in an amendment filed on February 9, 2001 to include the associated SEQ ID NOs. in the description of the figure, i.e., on page 32. A copy of the filed amendment is enclosed herein.

The specification is also objected to because of its recitation of "asparagene" on page 16, its recitation of "901C" on page 7, and its recitation of "26143" on page 8. Applicants have corrected these typographical errors.

The specification has been amended on page 56 to incorporate the structures of the exo motifs as taught in Derbyshire V et al., (1995), *Methods. Enzymol.* 262:363 (Exhibit A), which is incorporated by reference in its entirety as stated on page 33, line 5. Support for such amendments can be found in Derbyshire V et al., (1995), e.g., on page 383, first paragraph. No new matter is added.

In view of the above amendments in the specification, Applicant respectfully requests the objection on the specification be withdrawn.

Drawings

Figures 4-8 are being objected to because of alleged half-tone quality; Figures 7, 10-12, 14-15 are being objected to because of their margins; Figure 4 is objected to because it "should be labeled 4A-4C; Figure 13 is objected to because the shading should be removed.

Applicant submits that Figures 4-8 are resubmitted with better quality; the margins of Figures 7, 10-12 and 14-15 are fixed to comply with 37 C.F.R. §1.84(g); the shading in Figure

13 is removed. As to the label of Figure 4, Applicant submits that it is not clear as to how the Examiner wants the figure to be relabeled because the figure contains one single sequence. Applicant has relabeled Figure 5 as 5A-5C. A marked-up sheet showing the changes for Figure 5 is attached.

In view of the above amendment, Applicants respectfully request the objection on the figure be withdrawn.

Claim Objection

Claim 32 is objected to because of its recitation of "asparagene." Claim 32 is cancelled.

Claim Rejections

Claims 2, 3, 6-10, 12-47 and 85-88 are cancelled. Claims 89-127 are currently added. As the result of the present claim amendment, claims 1, 5 and 89-127 are pending. Claims 1 and 5 are allowed.

Claim Rejections—35 U.S.C. §112, First Paragraph, Lack of Enablement

Applicants have clarified the rejections with the Examiner in our interview on July 7, 2003, that claims currently rejected under 35 U.S.C. §112, First Paragraph for alleged lack of enablement are claims 2 and 88 only. The Examiner has further confirmed that the previous enablement rejections on claims 6-15 were withdrawn. Claims 2 and 88 are rejected on the same grounds as set forth for claims 6-15 in the previous Office Action of May 7, 2002, that is, the use of *Thermococcus* species JDF-3 is essential and that a deposit of the organism is required for claims 2 and 88.

Both claims 2 and 88 are cancelled, therefore making the enablement rejections moot, however, Applicants maintain the position that the deposit is not necessary because both claims 2 and 88 are limited by the sequence of SEQ ID NO: 2. In particular, claim 2 is dependent from claim 1, therefore further limits the allowed claim 1 in that the DNA polymerase having an

amino acid sequence presented in SEQ ID NO:2 is isolated from *Thermococcus* species JDF-3. Given the polynucleotide sequence of SEQ ID NO: 2 and the knowledge in the art of recombinant DNA and the teaching of the present specification, a skilled artisan can make and use the polymerase as recited in claims 2 and 88.

Claim Rejections—35 U.S.C. §112, First Paragraph, Lack of Written Description

Claims 6-10, 12-47, and 85-87 are rejected for lack of written description based on the same grounds as set forth in the previous Office Action of May 7, 2002.

The previous Office Action of May 7, 2002 states that "while Applicants have described a number of additional Family B polymerases and cited references which compare the sequences of many of these polymerases, Applicants have not described mutations which result in the desired polymerase properties in addition to those referred to in the previous Office Action." The Office Action further states that "the specification only provides the representative species encompassed by these claims, wherein said mutant polymerase is from *Thermococcus* species JDF-3 and the mutation is selected from the group consisting of mutations at residues: S345, P410, D141, E143, A485 and L408, of SEQ ID NO: 2." The Office Action continues:

> "While it is admitted that Applicants disclose a number of mutations, these are not representative of the genus of mutations claimed which encompasses any and all mutations of any Family B or *Thermococcus* species JDF-3 DNA polymerase which results in a decrease in 3' to 5' exonuclease activity or a reduction in discrimination against non-conventional nucleotides."

The Office Action concludes that "Given this lack of additional representative species encompassed by the claims, Applicants have failed to sufficiently describe the claimed invention in such full, clear, concise and exact terms that a skilled artisan would recognize applicants were in possession of the claimed invention."

In summary, the claims 6-10, 12-47, and 85-87 are rejected for their recitation of phrases "family B DNA polymerase" and/or "3' to 5' exonuclease deficient."

Applicants submit that claims 6-10, 12-47, and 85-87 are cancelled. New claims 89-127 are added to replace the cancelled claims. The new added claims do not recite the phrase "3' to 5' exonuclease deficient," claims 99-105. 101-102, and 124-125, however, do recite the phrase "family B DNA polymerase." The added claims 89-127 find support in the cancelled claims and throughout the specification, they satisfy the written description requirement as described below.

MPEP 2163 provides the following guidelines for the written description requirement under 35 U.S.C. §112, First Paragraph:

> "The analysis of whether the specification complies with the written description requirement calls for the examiner to compare the scope of the claim with the scope of the description to determine whether applicant has demonstrated *possession* of the claimed invention. *Such a review is conducted from the standpoint of one of skill in the art at the time the application was filed* (see, e.g., *Wang Labs. v. Toshiba Corp.*, 993 F.2d 858, 865, 26 USPQ2d 1767, 1774 (Fed. Cir. 1993)) and *should include a determination of the field of the invention and the level of skill and knowledge in the art.*" (*Emphasis added*)

Applicants submit claims 89-98, 108-123 and 126-127 are drawn to a DNA polymerase comprising a sequence of SEQ ID NO: 2 or a kit comprising such a DNA polymerase; claims 99-107 and 124-125 are drawn to a DNA polymerase comprising a sequence selected from the sequences as indicated by accession numbers listed in Table II or a kit comprising such a DNA polymerase. Claims 89-127 are further drawn to one or more amino acid mutations at specified amino acid positions or within specified regions recognized in the art. Applicants submit that, given the high level of skill in the art of recombinant DNA, the specification provides adequate support for the sequences recited in the claims 89-127 (i.e., SEQ ID NO: 2 and the sequences as indicated by accession numbers listed in Table II). Where certain claims are considered genus claims, e.g., for comprising one or more mutations within a specific region, the specification describes a representative number of species for each of the genus claims and provides adequate

support. Claims 89-127, therefore, satisfy the written description requirement under the above MPEP guideline.

First, the specification provides sufficient description for each of currently added claims with respect to its limitation to comprising a specific amino acid sequence.

MPEP 2163 provides that written description may be satisfied buy the description of "identifying characteristic," and that "[f]or some biomolecules, examples of identifying characteristics include a *sequence*, *structure*, binding affinity, binding specificity, molecular weight, and length...."

Applicants submit that currently added claims 89-98, 108-123 and 126-127 are drawn to an isolated recombinant DNA polymerase comprising a sequence of *SEQ ID NO: 2* and further comprising a mutation in specified region (e.g., exo motifs or Region II), or amino acid positions. Each of claims 89-98, 108-123 and 126-127 are therefore limited to comprising the sequence of SEQ ID NO: 2, with the exception that there are one or more mutations within the specified region (i.e., exo motifs or Region II) or at specific amino acid positions within the sequence of SEQ ID NO: 2.

The specification provide detailed teachings for SEQ ID NO: 2 (e.g., in Figure 2 and throughout the specification).

Currently added claims 99-105 and 124-125 are not limited to comprising SEQ ID NO: 2, however, they are limited to comprising a sequence selected from the sequences as indicated by accession numbers listed in Table II.

There are seventeen family B DNA polymerases listed in Table II (e.g., on pages 37-39). Each of the sequences in Table II is given its Genbank accession number, and each listed Genbank accession number represents only one family B DNA polymerase sequence. Claims 99-107 and 124-125 are limited to comprising one of the seventeen DNA polymerase sequences in Table II. Genbank accession number is well recognized in the art of recombinant DNA technology, one skilled in the art would have easily known what the sequence for a family B

DNA polymerase is as recited in claims 99-107 and 124-125 by using the accession numbers provided in Table II.

According to MPEP 2163:

> "Generally, there is an inverse correlation between the level of skill and knowledge in the art and the specificity of disclosure necessary to satisfy the written description requirement. *Information which is well known in the art need not be described in detail in the specification.* See, e.g., Hybritech, Inc. v. Monoclonal Antibodies, Inc., 802 F.2d 1367, 1379-80, 231 USPQ 81, 90 (Fed. Cir. 1986)....
>
> *If a skilled artisan would have understood the inventor to be in possession of the claimed invention at the time of filing, even if every nuance of the claims is not explicitly described in the specification, then the adequate description requirement is met.* See, e.g., *Vas-Cath*, 935 F.2d at 1563, 19 USPQ2d at 1116; Martin v. Johnson, 454 F.2d 746, 751, 172 USPQ 391, 395 (CCPA 1972) (stating "the description need not be in *ipsis verbis* [i.e., 'in the same words'] to be sufficient')." (*Emphasis added*)

Applicants submit that the level of skill in the art of recombinant DNA is high and that one skilled in the art would have understood that Applicants had possession of SEQ ID NO: 2 as recited in claims 89-98, 108-123 and 126-127 and the sequences as indicated by accession numbers listed in Table II as recited in claims 99-106 and 124-125 at the time of filing. Based on the above MPEP guideline, the teachings in the specification provide sufficient description with respect to SEQ ID NO: 2 and the sequences listed in Table II.

Second, the specification provides adequate description for each of claims 89-124 with respect to the specific regions (i.e., exo I, II, III, and Region II) in which mutations may occur.

As discussed above, all currently added claims are limited to having one or more mutations at specific amino acid positions and/or within four specific regions, i.e., exo I (DXE),

exo II ($NX_{2-3}FD$), exo III (YX_3D) motif and Region II within SEQ ID NO: 2 or a sequence selected from the group consisting of sequences as indicated by accession numbers listed in Table II.

MPEP 2163 provides:

> "The written description requirement for a claimed genus may be satisfied through sufficient description of *a representative number of species* by actual reduction to practice..., reduction to drawings..., or *by disclosure of relevant, identifying characteristics, i.e., structure or other physical and/or chemical properties, by functional characteristics coupled with a known or disclosed correlation between function and structure*, or by a combination of such identifying characteristics, sufficient to show the applicant was in possession of the claimed genus.... See *Eli Lilly*, 119 F.3d at 1568, 43 USPQ2d at 1406.
>
> A '*representative number of species*' means that *the species which are adequately described are representative of the entire genus*." (*Emphasis added*)

DNA polymerases comprising one or more mutations at specific amino acid positions or a specified region (e.g., exo I and Region II) within SEQ ID NO: 2, e.g., as recited in claims 89-98, 108-123 and 126-127, are described in the specification and are actually reduced to practice, for example, see 11-22, Figures 5-15. With respect to DNA polymerases comprising a Table II sequence and one or more mutations within the four specific regions (i.e., exo I, II, III or Region II), Applicants submit that the specification not only provides detailed description for the structure and function of the four regions, but it also provides a representative numbers of mutations within these regions.

With respect to exo motifs I, II, III, the amended specification teaches their structure requirements on page 56:

> "DNA polymerases lacking 3'-5' exonuclease (proofreading) activity are preferred for applications requiring nucleotide analog incorporation (e.g., DNA

sequencing) to prevent removal of nucleotide analogs after incorporation. The 3'-5' exonuclease activity associated with proofreading DNA polymerases can be reduced or abolished by mutagenesis. *Sequence comparisons have identified three conserved motifs (exo I (DXE), II (NX_{2-3}(F/Y)D), III (YX_3D)) in the 3'-5' exonuclease domain of DNA polymerases* (reviewed V. Derbyshire, J.K. Pinsonneault, and C.M. Joyce, Methods Enzymol. 262, 363 (1995)). Replacement of any of the conserved aspartic or glutamic acid residues with alanine has been shown to abolish the exonuclease activity of numerous DNA polymerases, including archaeal DNA polymerases such as Vent (H. Kong, R.B. Kucera, and W.E. Jack, J. Biol. Chem. 268, 1965 (1993)) and *Pfu* (Stratagene, unpublished). Conservative substitutions lead to reduced exonuclease activity, as shown for mutants of the archaeal 9° N-7 DNA polymerase (M.W. Southworth, H. Kong, R.B. Kucera, J. Ware, H. Jannasch, and F.B. Perler, Proc. Natl. Acad. Sci. 93, 5281 (1996))."

Thus, the specification as recited above specifically describes the structure of the exo motifs and their relationship to the function of exonuclease activity. Each of the exo motifs contains *only 3 to 6 amino acids*, among which 2-3 amino acids are conserved for family B DNA polymerases.

For claims drawn to DNA polymerase comprising SEQ ID NO: 2, the specification as recited above further teaches four mutants within exo I which reduce 3' to 5' Exo activity. The specification further provides description of the structure/function relationship between the conserved motifs and particular amino acid changes (i.e., replacement of glutamic acid or aspartic acid residues with alanine) one could make within those motifs (e.g., see above). Applicants have therefore provided an actual reduction to practice of four individual embodiments, plus functional characteristics (i.e., reduced Exo activity) coupled with a known and disclosed correlation between function and structure.

The specification also describes six different individual mutants of *Thermococcus* species JDF-3 Family B DNA polymerase that target a region correlated with 3' to 5' Exo activity in related polymerases (e.g., pages 11-22, and pages 57-58). These include D141A, D141N, D141S, D141T, D141E and E143A. The specification also describes the double mutant D141A + E143A (e.g., on page 58), for a total of 7 mutants within the three exo motifs. Three of the single mutants and the double mutant exhibited dramatically reduced 3' to 5' Exo activity. Thus, the specification provides an actual reduction to practice through four working examples of different mutants that fall within the claims reciting exo I, exo II or exo III motifs.

For claims 99-107 and 124-125, which are drawn to a sequence selected from the sequences as indicated by accession numbers listed in Table II, Applicants submit that one skilled in the art would recognize the exo I, II, III motifs in these Table II family B DNA polymerase sequences. The exo I, II, III motifs are well conserved among the family B DNA polymerases listed in Table II. One skilled in the art can easily take one sequence as indicated by an accession number listed in Table II and align it to another family B DNA polymerase sequence (e.g., SEQ ID NO: 2) to identify the corresponding exo I, II, and III motifs. Applicants submit herewith an amino acid sequence alignment of six different family B DNA polymerases taught in Table II (Exhibit B) to show that the three exo motifs in other family B DNA polymerases align with those of SEQ ID NO: 2 (i.e., the JDF-3 sequence) and can be easily identified by one skilled in the art.

According to MPEP 2163 guideline "when there is substantial variation within the genus, one must describe a sufficient variety of species to reflect the variation within the genus. On the other hand, *there may be situations where one species adequately supports a genus*.... What constitutes *a "representative number" is an inverse function of the skill and knowledge in the art.* Satisfactory disclosure of *a "representative number" depends on whether one of skill in the art would recognize that the applicant was in possession of the necessary common attributes or features of the elements possessed by the members of the genus in view of the species disclosed*.... Description of a representative number of species *does not require the description*

to be of such specificity that it would provide individual support for each species that the genus embraces."

Because of the high level of skill in the art of recombinant DNA, and because the specification provides a number of mutants within the exo motifs (i.e., exo I, exo II, and exo III), along with a description of conserved structural regions involved in such functional activity and specific mutations one should make in those regions in order to reduce 3' to 5' Exo activity, Applicants submit that the written description requirement is satisfied with regard to the exo motifs.

With respect to Region II (DXXSLYPSII), Applicants submit that the specification satisfy the written description requirement under 35 U.S.C. §112, first paragraph.

When the claims (i.e., 89-98, 108-123 and 126-127) are drawn to a DNA polymerase comprising SEQ ID NO: 2, this Region II consensus sequence is present at amino acids 404 to 413 of SEQ ID NO: 2. Tables V and VI on pages 74 and 75 of the specification describe the isolation and testing of 12 DNA polymerase mutants with mutations within this consensus region of SEQ ID NO: 2. Specifically, Table VI describes the isolation of 7 mutants with a primary mutation at L408 (3 are L408H, 4 are L408F) and 4 mutants with a primary mutation at P410 (all P410 L). Table V describes an additional P410L mutant (mutant p11). Thus, the specification provides 12 mutants, representing 3 different mutations within the Region II consensus sequence of SEQ ID NO: 2 as recited in claims 91-98, 110-121 and 126. Each of these mutants has a reduced discrimination against non-conventional nucleotides.

In addition to the description of the mutants in Tables V and VI, the specification provides additional description of numerous double mutants (a total of over 60 mutants are described in the specification). Ten such double mutants comprise at least one mutation within the Region II consensus region within the sequence of SEQ ID NO: 2. These include: L408H + A485T, L408F + A485T, and P410L + A485T described on page 15, lines 14-22; P410H + S345P and P410L + S345P on page 16, lines 1-3; L408H + V437 and L408H + L478 on page 17, lines 4-7; and A485T + Y409V, L408 mutation + Y409V and P410 mutation + Y409V. In all,

five different amino acid mutations are described within the Region II consensus sequence recited in claim 10: L408H, L408F, P410L, P410H and P409V.

The specification further addresses the function of the Region II consensus structure of SEQ ID NO: 2 at page 52, where it states:

> "The domains of relevance in 17 of the 40 purified mutants were sequenced. Most randomly mutated clones contained more than one mutation in the regions sequenced but all mutants contained mutations at one of three sites. Mutations predicted to confer an enhanced ddNTP uptake phenotype were introduced into the progenitor exonuclease deficient DNA polymerase sequence by site-directed mutagenesis to eliminate ancillary mutations which were not expected to contribute to the improved dideoxynucleotide uptake phenotype.
>
> Sixteen of the seventeen JDF-3 DNA polymerase mutations were found in Region II (motif A) on either side of the tyrosine in the consensus sequence 404 DxxSLYPSII 413. These mutations consisted of DFRSLY**L**SII (P410L), DFRS**H**YPSII (L408H) and DFRS**F**YPSII (L408F)." (page 52, lines 13-23; **emphasis added)**

This passage points out that the mutations expected to have an effect on improved dideoxynucleotide uptake were centered in the Region II consensus sequence of SEQ ID NO: 2 as recited in claims 91-98, 110-123 and 126-127. This description provides a further structure/function correlation between the Region II consensus DXXSLYPSII and nucleotide discrimination.

With respect to claims (e.g., 99-107 and 124-125) drawn to a DNA polymerase comprising a sequence selected from the sequences as indicated by accession numbers listed in Table II, Applicants submit that the specification provides description of other Family B DNA polymerases mutated in the Region II consensus and the effects of these mutations on discrimination against non-conventional nucleotides at page 7, line 20 to page 9, line 19. In

particular, in reference to Y409 of the consensus sequence, the specification states "Mutagenesis studies done in Family B DNA polymerases also implicate the containing the analogous Y in Region II in dNTP incorporation and ribose selectivity" (see p. 8, lines 11-12). The specification then continues to describe additional Region II mutations in other Family B DNA polymerases including the human DNA polymerase α (mutation at the site corresponding to Y409), bacteriophage T4 DNA polymerase (two mutations at the site corresponding to L408), bacteriophage ϕ29 DNA polymerase (mutations at the sites corresponding to L408 and P410), and the archaeal Family B DNA polymerase from *Thermococcus litoralis* (VENT; three mutations at a site corresponding to Y409) (see p. 8, line 12 to page 9, line 19). These mutants, the mutations of which each fall within the Region II consensus DXXSLYPSII, are discussed with regard to the impact of the mutations on nucleotide discrimination. Thus, the specification makes it clear that Region II structure, which is defined by the consensus DXXSLYPSII, is important in the function of nucleotide discrimination by Family B DNA polymerases.

In addition, Applicants submit that Region II sequence is conserved among the family B DNA polymerases listed in Table II. One skilled in the art can easily take a family B DNA polymerase sequence as indicated by an accession number listed in Table II and align it to another family B DNA polymerase (e.g., SEQ ID NO: 2) to identify the corresponding Region II sequence in this family B DNA polymerases. Applicants submit herewith an amino acid sequence alignment of six different family B DNA polymerases (Exhibit C) to show that Region II in family B DNA polymerases aligns with that of SEQ ID NO: 2 (i.e., JDF-3 sequence) and can be easily identified by one skilled in the art.

Given the description in the specification of the structure/function relationship between the Region II consensus sequence DXXSLYPSII and reduced discrimination against non-conventional nucleotides, *and* the numerous (*at least 22*, given 12 mutants in Tables V and VI and 10 double mutants described on pages 15-17) single and double mutants described in the specification that have mutations falling within this consensus sequence of Region II, Applicants submit that a representative number of reduced discrimination mutants in the region recited in claims 99-107 and 124-125 have been described to show that Applicants were in possession of

the claimed genus. That is, applicants have adequately described a representative number of species within the claimed genus of isolated recombinant DNA polymerases comprising SEQ ID NO: 2 or a sequence selected from the sequences as indicated by accession numbers listed in Table II and further comprising one or more mutations within the Region II consensus sequence DXXSLYPSII.

In view of the above, Applicants submit that the currently added claims 89-127 satisfy the written description requirement under 35 U.S.C. §112, First Paragraph.

Claim Rejections -- 35 U.S.C. §112, Second Paragraph

Claim 85 is rejected for reciting the limitation "said mutation in Region II" due to insufficient antecedent basis.

Claim 85 is cancelled.

Claim Rejections -- 35 U.S.C. §102(a)

Claims 6, 10, 14, 15, and 87 are rejected under 35 U.S.C. §102(a) as being anticipated by Gardner et al. The Examiner states that Gardner teaches a Y412V Vent DNA polymerase variant that incorporates ribonucleotides at least 200 fold more efficiently than the wild-type enzyme.

Applicants submit that claims 6, 10, 14, 15 and 87 are cancelled.

The currently added claims 89-98, 108-123 and 126-127 are drawn to a DNA polymerase comprising a sequence of SEQ ID NO: 2 and one or more mutations at specific amino acid positions or specific regions, including a DNA polymerase comprising a sequence of SEQ ID NO: 2 and a Y412V mutation. Y412V is a mutation within Region II of Vent DNA polymerase (e.g., see Exhibit B). Gardner does not teach a DNA polymerase comprising a sequence of SEQ ID NO: 2 and a Y412V mutation. Therefore claims 89-98, 108-123 and 126-127 are not anticipated by Gardner.

The currently added claims 99-100, 103-107 and 124-125 are drawn to a family B DNA polymerase comprising a sequence selected from the sequences as indicated by accession numbers listed in Table II, including vent DNA polymerase, and may further comprising one or more mutations at specific amino acid positions or specific regions. The claims as written only recites a mutation at the leucine and/or proline positions in the Region II consensus sequence, therefore, do not include a mutation at Y412 as taught in Gardner. Therefore, Gardner does not anticipate claims 99-100, 103-107 and 124-125.

The currently added claims 101-102 are also drawn to a family B DNA polymerase comprising a sequence selected from the sequences as indicated by accession numbers listed in Table II, including vent DNA polymerase, and may further comprising one or more mutations at specific amino acid positions or specific regions. The claims as written recites a mutation in the region II consensus sequence, including Y412, however, the claims also recite an additional mutation at an amino acid corresponding to A485 of SEQ ID NO: 2. Gardner does not teach a vent DNA polymerase comprising a mutation in Region II and a mutation at the position corresponding to A485 of SEQ ID NO: 2, therefore, it does not anticipate claims 101-102.

Claim Rejections -- 35 U.S.C. §102(b)

Claims 6, 10, 14, 15 and 87 are rejected under 35 U.S.C. §102(b) as being anticipated by Dong-et al. The Examiner states that Dong et al. teaches a mutant of the family B DNA polymerase, i.e., human DNA polymerase α, comprising Y865S and Y865F, where the mutant is 3' to 5' exonuclease deficient, and has a reduced discrimination against non-conventional nucleotides.

Applicants submit that claims 6, 10, 14, 15 and 87 are cancelled.

The currently added claims 89-98, 108-123 and 126-127 are drawn to a DNA polymerase comprising a sequence of SEQ ID NO: 2 and one or more mutations at specific amino acid positions or specific regions. Dong et al. does not teach a DNA polymerase comprising SEQ ID NO: 2, therefore, does not anticipate claims 89-98, 108-123 and 126-127.

The currently added claims 99-107 and 124-125 are drawn to a family B DNA polymerase comprising a sequence selected from the sequences as indicated by accession numbers listed in Table II, not including human DNA polymerase α. Dong et al. does not teach a sequence selected from the sequences as indicated by accession numbers listed in Table II, therefore, it does not anticipate the currently added claims 99-107 and 124-125.

Applicants submit that all claims (i.e., claims 1, 5, 89-127) are allowable as currently written and respectfully requests favorable action by the Examiner. If the Examiner believes that a telephone conversation with Applicant's attorney/agent would expedite prosecution of this application, the Examiner is cordially invited to call the undersigned attorney/agent of record.

Respectfully submitted,

Date: July 25, 2003

Name: Kathleen M. Williams
Registration No.: 34,380
Customer No.: 27495
Palmer & Dodge LLP
111 Huntington Avenue
Boston, MA 02199-7613
Tel: 617-239-0100

Visualization of Areas of Clearing

ning, the stack is removed and the gel
ion/reaction buffer, typically in a small
buffer for M-MuLV RT consists of 50
, 50 m*M* NaCl, and 2 m*M* $MnCl_2$. The
s above with the substitution of 10 m*M*
ould be as fresh as possible and added
nt oxidation of the $MnCl_2$. If the buffer
idized and no activity will be detected.
pH of the Tris–HCl from 8.0 to 7.5–7.2
l should shake for about 48 hr at either
the buffer approximately every 12 hr.
portant, helping remove residual SDS
m the gel. After 48 hr, the wet gel is
ure to see if enough renaturation and
s placed in a sealable plastic bag, which
c wrap. A wet gel typically only needs
et an adequate exposure. If not enough
zones of clearing are detected), we let
e or two more days. Finally the gel is
film by standard methods. An example
ig. 3.

NA:DNA and RNA:RNA substrates,
ever, we observed that these substrates
ring the electrophoresis or the renatu-

veral RNA:DNA and RNA:RNA
coli, and it may be necessary to experi-
ercentages of acrylamide to resolve the
ting RNases H. *Escherichia coli* RNase
D) migrate far from RT, and serve as
Nase III-deficient strains of *E. coli* can
s if necessary.[4,16,19]

of Public Health Service and CA30488 to SPG and
estigator of the Howard Hughes Medical Institute.

[28] Structure–Function Analysis of 3′ → 5′-Exonuclease of DNA Polymerases

By VICTORIA DERBYSHIRE, JULIA K. PINSONNEAULT, and CATHERINE M. JOYCE

Introduction

The 3′ → 5′-exonuclease activity of DNA polymerases acts in opposition to the polymerase activity and serves as a proofreader, by removing polymerase errors.[1] This activity is present in the majority of DNA-dependent DNA polymerases but absent in the reverse transcriptase (RT) family. The 3′ → 5′-exonuclease is usually part of the same polypeptide chain as the DNA polymerase; an exception is the multisubunit DNA polymerase III of *Escherichia coli,* where the editing function is present on a separate subunit (ε) within the core polymerase.[2] In the structure of the Klenow fragment of DNA polymerase I, the 3′ → 5′-exonuclease is located on a discrete structural domain,[3] and it seems likely that other DNA polymerases are arranged in a similar modular fashion. The preferred substrate for the exonuclease is single-stranded DNA, and a variety of data are consistent with the idea that the primer terminus of a duplex DNA substrate is bound as a "frayed" or single-stranded end at the exonuclease active site.[4] Of all the reactions catalyzed by DNA polymerases, the 3′ → 5′-exonuclease is probably the best understood at a mechanistic level, thanks to the crystallographic data obtained with the Klenow fragment of *E. coli* DNA polymerase I. By studying cocrystals containing either the substrate DNA or the product (dNMP) at the 3′ → 5′-exonuclease site, it has been possible to identify the side chains that interact with substrate or product and form the active site[3,5,6] (Fig. 1). The structural data provided the basis for a detailed mutagenesis study of the roles of these side chains in the exonuclease reaction.[7,8] The mutational study contributed to the crystallographic characterization

[1] A. Kornberg and T. A. Baker, "DNA Replication" p. 113. Freeman, San Francisco, 1992.
[2] R. H. Scheuermann and H. Echols, *Proc. Natl. Acad. Sci. USA* **81,** 7747 (1984).
[3] D. L. Ollis, P. Brick, R. Hamlin, N. G. Xuong, and T. A. Steitz, *Nature* **313,** 762 (1985).
[4] L. S. Beese, V. Derbyshire, and T. A. Steitz, *Science* **260,** 352 (1993).
[5] P. S. Freemont, J. M. Friedman, L. S. Beese, M. R. Sanderson, and T. A. Steitz, *Proc. Natl. Acad. Sci. USA* **85,** 8924 (1988).
[6] L. Beese and T. A. Steitz, *EMBO J.* **10,** 25 (1991).
[7] V. Derbyshire, P. S. Freemont, M. R. Sanderson, L. Beese, J. M. Friedman, C. M. Joyce, and T. A. Steitz, *Science* **240,** 199 (1988).
[8] V. Derbyshire, N. D. F. Grindley, and C. M. Joyce, *EMBO J.* **10,** 17 (1991).

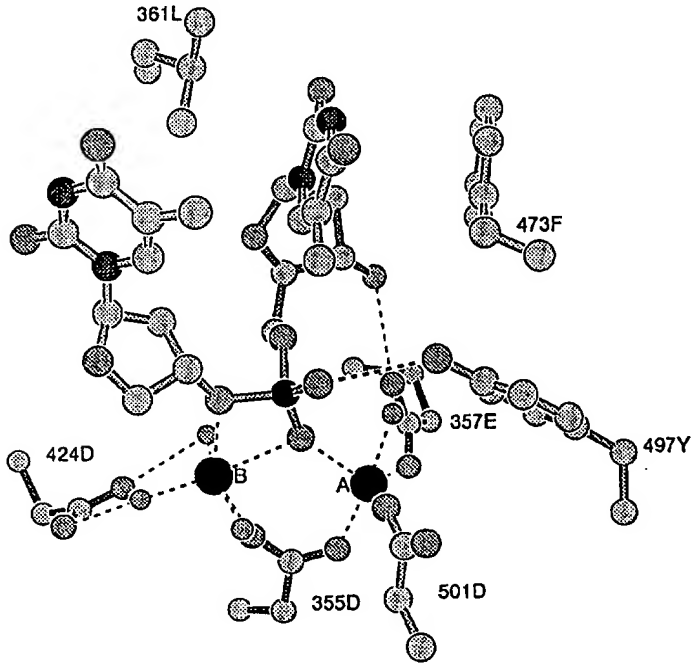


FIG. 1. Structure of the 3′ →5 ′-exonuclease active site containing a bound dinucleotide. The catalytically essential metal ions, A and B, are shown as large black balls. Other atoms are represented as smaller balls, with phosphorus black, and carbon, oxygen, and nitrogen represented by increasingly darker shades of gray. Water molecules are shown as smaller gray spheres. Reproduced, with permission from Oxford University Press, from L. Beese and T. A. Steitz, *EMBO J.* **10,** 25 (1991).

of two divalent metal ion sites at the 3′ → 5′-exonuclease site and showed the importance of the metal ions in the reaction.[7,8] Based on the structural data, a reaction mechanism was proposed involving catalysis by the two metal ions.[5,6]

It has become clear from protein sequence alignments that all polymerases with an editing function possess the same group of crucial active site residues identified in Klenow fragment, although the surrounding protein sequence may be very dissimilar.[9–12] Moreover, preliminary crystallographic

[9] A. Bernad, L. Blanco, J. M. Lázaro, G. Martin, and M. Salas, *Cell* **59,** 219 (1989).
[10] D. K. Braithwaite and J. Ito, *Nucleic Acids Res.* **21,** 787 (1993).
[11] A. Morrison, J. B. Bell, T. A. Kunkel, and A. Sugino, *Proc. Natl. Acad. Sci. USA* **88,** 947 (1991).
[12] L. Blanco, A. Bernad, and M. Salas, *Gene* **112,** 139 (1992).

data for the N-terminal 45-kDa dom
the prediction of a similar active site g
provided the rationale for mutagenes
variety of DNA polymerases, which se
of a conserved 3′ → 5′-exonuclease ac
deficient enzymes for a variety of app

In this chapter, we describe first th
ducted on the Klenow fragment 3′ → 5
similar studies on other DNA polymer
ies, and the resulting mechanistic de
where.[14]

Mutagenesis of 3′ → 5′-Exonuclease

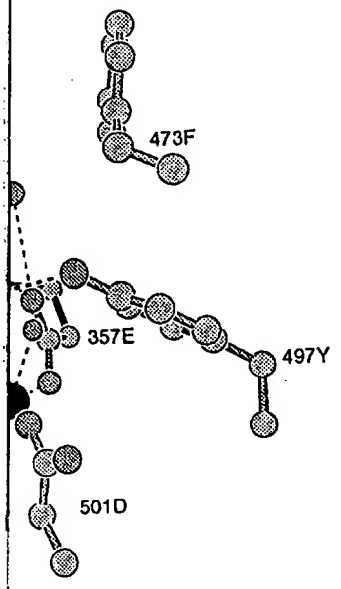
Choice of Mutations

Residues were chosen for mutagene
lographic data for the complexes of K
product at the exonuclease active site[3]
be divided into three fairly distinct g
anchoring the two divalent ions, those
the terminal phosphodiester bond tha
contact the upstream portion of the D
site of catalysis (Table I). In every cas
to alanine was studied; this is likely t
simple removal of the side chain, since al
disruptive than the alternative, glycine. I
tive changes were made; most of these w
of a similar size to the wild-type residue
properties. The rationale for such change
the experimental results below. In additi
made to test the tolerance for altered g
and the L361M mutation was made to inc
of the ε subunit of DNA polymerase II

Construction of Mutations

Site-directed mutagenesis was carrie
dures.[15] Synthetic oligonucleotides enc
used as primers on uracil-containing M1

[13] J. Wang, P. Yu, W. H. Konigsberg, and T. A. S
[14] C. M. Joyce and T. A. Steitz, *Ann. Rev. Bioche*
[15] T. A. Kunkel, J. D. Roberts, and R. A. Zakour



active site containing a bound dinucleotide. are shown as large black balls. Other atoms rus black, and carbon, oxygen, and nitrogen gray. Water molecules are shown as smaller n Oxford University Press, from L. Beese and

3′ → 5′-exonuclease site and showed ie reaction.[7,8] Based on the structural posed involving catalysis by the two

equence alignments that all polymer- the same group of crucial active site nt, although the surrounding protein Moreover, preliminary crystallographic

rtin, and M. Salas, *Cell* **59,** 219 (1989).
Res. **21,** 787 (1993).
d A. Sugino, *Proc. Natl. Acad. Sci. USA* **88,**

112, 139 (1992).

data for the N-terminal 45-kDa domain of T4 DNA polymerase confirm the prediction of a similar active site geometry.[13] The sequence alignments provided the rationale for mutagenesis of the 3′ → 5′-exonuclease in a variety of DNA polymerases, which served both to confirm the hypothesis of a conserved 3′ → 5′-exonuclease active site and to provide exonuclease-deficient enzymes for a variety of applications.

In this chapter, we describe first the mutational studies that were conducted on the Klenow fragment 3′ → 5′-exonuclease, and then summarize similar studies on other DNA polymerases. These structure–function studies, and the resulting mechanistic deductions, have been reviewed elsewhere.[14]

Mutagenesis of 3′ → 5′-Exonuclease of Klenow Fragment

Choice of Mutations

Residues were chosen for mutagenesis based on high-resolution crystallographic data for the complexes of Klenow fragment with substrate or product at the exonuclease active site[3,5,6] (Fig. 1). Active-site residues can be divided into three fairly distinct groups: those that serve as ligands anchoring the two divalent ions, those that contact the substrate around the terminal phosphodiester bond that is to be cleaved, and those that contact the upstream portion of the DNA chain, more remote from the site of catalysis (Table I). In every case a mutation of the target residue to alanine was studied; this is likely to be the best approximation to a simple removal of the side chain, since alanine is less likely to be structurally disruptive than the alternative, glycine. In many cases, other more conservative changes were made; most of these were substitutions giving side chains of a similar size to the wild-type residue but with altered hydrogen-bonding properties. The rationale for such changes should be apparent as we describe the experimental results below. In addition, changes from Asp to Glu were made to test the tolerance for altered geometry at positions 424 and 501, and the L361M mutation was made to increase the similarity to the sequence of the ε subunit of DNA polymerase III.

Construction of Mutations

Site-directed mutagenesis was carried out following established procedures.[15] Synthetic oligonucleotides encoding the desired mutations were used as primers on uracil-containing M13 templates containing appropriate

[13] J. Wang, P. Yu, W. H. Konigsberg, and T. A. Steitz, unpublished observations (1993).
[14] C. M. Joyce and T. A. Steitz, *Ann. Rev. Biochem.* **63,** 777 (1994).
[15] T. A. Kunkel, J. D. Roberts, and R. A. Zakour, *Methods Enzymol.* **154,** 367 (1987).

TABLE I
AMINO ACID RESIDUES AT 3′ → 5′-EXONUCLEASE ACTIVE SITE OF KLENOW FRAGMENT

Residue	Observed contact[a]	Mutations[b]
Ligands to the metal ions		
Asp-355	Shared ligand to metal A and metal B	D355A, D355N
Glu-357	Ligand to metal A (also see below)	E357A, E357Q
Asp-424	Ligand to metal B (via bridging H_2O molecules)	D424A, D424E, D424N
Asp-501	Ligand to metal A	D501A, D501E, D501N
Substrate contacts at 3′ terminus		
Glu-357	Hydrogen-bonded to 3′-OH and may also orient attacking H_2O	E357A, E357Q
Leu-361	Inserted between terminal two bases	L361A, L361M
Phe-473	Stacks with 3′-terminal base	F473A
Tyr-497	Hydrogen-bonded to terminal phosphodiester bond and may also orient attacking H_2O	Y497A, Y497F
Upstream contacts with the substrate sugar–phosphate backbone[c]		
Gln-419	Interacts with penultimate phosphodiester bond	Q419A, Q419E
Arg-455	Ion-pair interaction with third phosphodiester bond from terminus	R455A

[a] From L. Beese and T. A. Steitz, *EMBO J.* **10,** 25 (1991).
[b] Mutations are abbreviated using the following convention: The residue number from the DNA polymerase I sequence [C. M. Joyce, W. S. Kelley, and N. D. F. Grindley, *J. Biol. Chem.* **257,** 1958 (1982)] is preceded by the symbol (in the one-letter code) for the wild-type amino acid and followed by the symbol for the mutant amino acid. Thus D355A denotes a mutation from Asp to Ala at position 355.
[c] Additional contacts with the nucleotide bases and deoxyribose positions are described by Beese and Steitz (see footnote *a*).

regions of the *polA* gene, chosen so as to facilitate subsequent cloning of the mutations into a Klenow fragment expression plasmid.[8] M13 isolates carrying the desired mutation were obtained typically at frequencies of 10 to 50% and were identified by direct sequencing of randomly chosen clones. Before subcloning into the expression plasmid, the region to be subcloned (300 to 500 bp, see below) was sequenced in its entirety to check that no additional mutations were present.

In a large-scale mutagenesis study such as the present one, it is extremely helpful to have a series of unique restriction sites that divides the target gene into modules or cassettes of a convenient size for cloning and sequencing. The 3′ → 5′-exonuclease region of Klenow fragment can be divided into two modules, defined by three unique restriction sites: the *Bst*XI site

in λ control sequences 70 bp upstream c
start (codon 324 of the *polA* structur
402–403, and the *Sac*I site at codons 5
due 402 were therefore retrieved from
*Xho*I fragment, and those beyond re
fragment. The relevant fragment was i
plasmid, in which Klenow fragment i
leftward promoter (P_L) of phage λ.[17,1
tated by the use of deletion derivatives
pair of unique sites (*Bst*XI and *Xho*I,
short (~10 bp) adaptor containing a *B*
of this cloning strategy was that there
type information instead of the desired
ing region was absent from the recipie
were that the cloning efficiency could
ligation mixture with *Bam*HI to reduce
and that plasmids containing the desired
tiated from the starting plasmid by the
restriction fragment. The overproducer
were obtained and characterized in stra
the *recA*⁻ host DH1.[19] We chose a *recA*
genetic exchange between the mutant
and the wild-type chromosomal *polA*
many of the commonly used cloning ve
study.

Because this study required the cloni
ment of time in constructing appropriate
tated fragments was clearly worthwhile; i
were to be studied, the benefit of such an
In the Klenow fragment system, the prese
site within the region under study was al
easy construction of some double mutatio
likewise found the "cassette" approach t
polymerase region of Klenow fragment,
also necessary to create, by mutation, t
modules within the coding sequence.[17]

[16] C. M. Joyce, W. S. Kelley, and N. D. F. Grindle
[17] A. H. Polesky, T. A. Steitz, N. D. F. Grindley
14,579 (1990).
[18] C. M. Joyce and V. Derbyshire, this volume, [1].
[19] D. Hanahan, *J. Mol. Biol.* **166,** 557 (1983).

E I
EASE ACTIVE SITE OF KLENOW FRAGMENT

contact[a]	Mutations[b]
A and metal B	D355A, D355N
o see below)	E357A, E357Q
bridging H_2O	D424A, D424E, D424N
	D501A, D501E, D501N
'-OH and may also	E357A, E357Q
p	
inal two bases	L361A, L361M
l base	F473A
erminal	Y497A, Y497F
l and may also orient	
ate backbone[c]	
nate phosphodiester	Q419A, Q419E
th third	R455A
l from terminus	

25 (1991).
convention: The residue number from the DNA
ey, and N. D. F. Grindley, *J. Biol. Chem.* **257,** 1958
ter code) for the wild-type amino acid and followed
is D355A denotes a mutation from Asp to Ala at

and deoxyribose positions are described by Beese

as to facilitate subsequent cloning of
nt expression plasmid.[8] M13 isolates
btained typically at frequencies of 10
equencing of randomly chosen clones.
plasmid, the region to be subcloned
nced in its entirety to check that no

such as the present one, it is extremely
striction sites that divides the target
nvenient size for cloning and sequenc-
of Klenow fragment can be divided
nique restriction sites: the *Bst*XI site

in λ control sequences 70 bp upstream of the Klenow fragment translational start (codon 324 of the *polA* structural gene[16]), the *Xho*I site at codons 402–403, and the *Sac*I site at codons 558–559. Mutations upstream of residue 402 were therefore retrieved from the M13 clone on a 302 bp *Bst*XI–*Xho*I fragment, and those beyond residue 403 on a 470 bp *Xho*I–*Sac*I fragment. The relevant fragment was inserted into the pCJ122 expression plasmid, in which Klenow fragment is under the control of the strong leftward promoter (P_L) of phage λ.[17,18] The cloning procedure was facilitated by the use of deletion derivatives of pCJ122 in which the appropriate pair of unique sites (*Bst*XI and *Xho*I, or *Xho*I and *Sac*I) was joined by a short (~10 bp) adaptor containing a *Bam*HI site. The primary advantage of this cloning strategy was that there was no danger of recovering wild-type information instead of the desired mutation because the corresponding region was absent from the recipient plasmid. Additional advantages were that the cloning efficiency could be increased by digestion of the ligation mixture with *Bam*HI to reduce the background of starting plasmid, and that plasmids containing the desired mutations could easily be differentiated from the starting plasmid by the increase in size of an appropriate restriction fragment. The overproducer plasmids for the mutant proteins were obtained and characterized in strain CJ388, a wild-type λ lysogen of the *recA*⁻ host DH1.[19] We chose a *recA*⁻ host to minimize the chances for genetic exchange between the mutant *polA* information on the plasmid and the wild-type chromosomal *polA* copy, needed for maintenance of many of the commonly used cloning vectors, including those used in this study.

Because this study required the cloning of many mutations, the investment of time in constructing appropriate "recipient plasmids" for the mutated fragments was clearly worthwhile; if only a small number of mutations were to be studied, the benefit of such an approach would be questionable. In the Klenow fragment system, the presence of the unique *Xho*I restriction site within the region under study was also beneficial in that it allowed the easy construction of some double mutations (e.g., E357A,D501N). We have likewise found the "cassette" approach to be valuable in our studies of the polymerase region of Klenow fragment, although in this instance it was also necessary to create, by mutation, the unique sites that defined the modules within the coding sequence.[17]

[16] C. M. Joyce, W. S. Kelley, and N. D. F. Grindley, *J. Biol. Chem.* **257,** 1958 (1982).
[17] A. H. Polesky, T. A. Steitz, N. D. F. Grindley, and C. M. Joyce, *J. Biol. Chem.* **265,** 14,579 (1990).
[18] C. M. Joyce and V. Derbyshire, this volume, [1].
[19] D. Hanahan, *J. Mol. Biol.* **166,** 557 (1983).

Overproduction and Purification of Mutant Proteins

Detailed procedures are given elsewhere in this volume.[18] Klenow fragment derivatives with mutations at the exonuclease site were overproduced by heat induction of strains carrying a temperature-sensitive λ repressor. Because the overproduction vector, carrying the mutated copy of the gene, requires host DNA polymerase I functions for its replication, two potential problems exist: (1) Contamination of a mutant Klenow fragment with wild type (derived by endogenous proteolysis of the host *polA* gene product) could give the appearance of exonuclease activity in a mutant protein that was, in reality, exonuclease deficient. (2) Recombination between mutant *polA* information on the expression plasmid and the wild-type chromosomal copy could convert the overproducer plasmid to wild type in a subpopulation of the culture. We circumvented both problems completely in the overproduction of the D355A,E357A double mutant protein by constructing an overproduction strain, CJ375, in which all the *polA* information is derived from the same exonuclease-deficient allele (this was possible because the D355A,E357A mutation does not affect plasmid replication).[20] Because it was impractical to construct a host strain of this type for every mutant protein being studied, other mutant proteins were expressed in a *recA* host strain, CJ376,[18] to reduce the likelihood of genetic exchange between mutant and wild-type information. As a further precaution against genetic exchange, overproducer strains were not stored as such; instead, cells for induction were grown from a fresh transformant in every case. The use of the CJ376 host strain addresses the problem of recombination, but does not deal with the possibility of low-level contamination with wild-type Klenow fragment.[21] Since the D355A,E357A overproducer system provides a more satisfactory solution to both problems, it is the D355A,E357A mutant protein that is available commercially as an exonuclease-deficient Klenow fragment, even though there are several other mutant proteins having similarly low levels of exonuclease activity (see Table II).

The mutant proteins were purified using the Pharmacia FPLC system.[18] We have found the superior fractionation obtained by FPLC to be particularly important for removing trace contaminants of cellular nucleases that might otherwise confuse the analysis of the mutant proteins. We avoid

[20] C. M. Joyce, unpublished observations (1987).

[21] The quantitative effect of both problems seems likely to be small, as shown by comparing the exonuclease activities of D355A,E357A (prepared from a homogenotized background), D424A (prepared from a *polA(D355A,E357A) recA*$^+$ background),[7] and D424N (prepared from a *polA*$^+$ *recA*$^-$ background).[8] The values (from Table II) are, respectively, 1.4×10^{-5}, 1.3×10^{-5}, and 2.5×10^{-5} of wild type.

TABL

ENZYMATIC ACTIVITY OF WILD-TYPE AND MU

Protein	Polymerase[b] activity
Wild type	1
Mutations affecting metal ligands	
D355A	1.4
D355N	0.7
E357A	1.3
E357Q	1.2
D424A	1.0
D424E	1.2
D424N	1.3
D501A	1.3
D501E	1.1
D501N	1.1
D355A, E357A	1.0
E357A, D501N	0.8
Mutations affecting contacts to terminal nucleot	
E357A[g]	1.3
E357Q[g]	1.2
L361A	1.0
L361M	1.2
F473A	1.1
Y497A	1.0
Y497F	1.2
Q419A	1.1
Q419E	1.5
R455A	0.8

[a] All values are the average of several determ the original reports.[7,8]

[b] Assayed on poly[d(AT)] template.[23] Values ar

[c] Relative to wild type (defined as 100 in both rates for wild-type Klenow fragment in the t and assay methods were those described in lower limit of the assay using double-strand of the single-stranded DNA assay so that it having very low activity.

[d] These assays were not necessarily under V_m rates at different substrate concentrations.

[e] This value was calculated from the rate of de indicated in the text, this assay is less sensiti For comparison, the D355A protein, assaye exonuclease rate, implying that the lower lim wild type as 100).

[f] No reaction detectable after a 300-min incuba

[g] Glu-357 is placed in both groups since the ca as a metal ligand and as a substrate-binding r

utant Proteins

where in this volume.[18] Klenow frag-
exonuclease site were overproduced
temperature-sensitive λ repressor.
rrying the mutated copy of the gene,
ions for its replication, two potential
mutant Klenow fragment with wild
ysis of the host *polA* gene product)
ase activity in a mutant protein that
(2) Recombination between mutant
smid and the wild-type chromosomal
plasmid to wild type in a subpopula-
d both problems completely in the
A double mutant protein by con-
375, in which all the *polA* information
se-deficient allele (this was possible
does not affect plasmid replication).[20]
ct a host strain of this type for every
mutant proteins were expressed in a
the likelihood of genetic exchange
ation. As a further precaution against
ins were not stored as such; instead,
a fresh transformant in every case.
resses the problem of recombination,
of low-level contamination with wild-
D355A,E357A overproducer system
tion to both problems, it is the
is available commercially as an exo-
, even though there are several other
w levels of exonuclease activity (see

d using the Pharmacia FPLC system.[18]
ation obtained by FPLC to be particu-
contaminants of cellular nucleases that
sis of the mutant proteins. We avoid

987).
seems likely to be small, as shown by comparing
A (prepared from a homogenotized background),
57A) *recA*⁺ background),[7] and D424N (prepared
lues (from Table II) are, respectively, 1.4×10^{-5},

TABLE II
ENZYMATIC ACTIVITY OF WILD-TYPE AND MUTANT DERIVATIVES OF KLENOW FRAGMENT[a]

Protein	Polymerase[b] activity	Ratio of exonuclease to polymerase activity[c]	
		Double-stranded DNA	Single-stranded DNA
Wild type	1	100	100
Mutations affecting metal ligands			
D355A	1.4	0.0083[d]	
D355N	0.7	≤0.01[e]	
E357A	1.3	0.18	
E357Q	1.2	0.03[d]	
D424A	1.0	0.0013[d]	
D424E	1.2	4.0	8.3
D424N	1.3	0.0025[d]	
D501A	1.3	0.0075[d]	
D501E	1.1	0.56	
D501N	1.1	50	
D355A, E357A	1.0	0.0014[d]	
E357A, D501N	0.8	≤0.002[f]	
Mutations affecting contacts to terminal nucleotide			
E357A[g]	1.3	0.18	
E357Q[g]	1.2	0.03[d]	
L361A	1.0	4.0	37
L361M	1.2	8.3	
F473A	1.1	0.03[d]	
Y497A	1.0	5.6	2.9[d]
Y497F	1.2	4.3	1.6
Q419A	1.1	23	20
Q419E	1.5	0.1[d]	
R455A	0.8	36	84

[a] All values are the average of several determinations; standard deviations are given in the original reports.[7,8]

[b] Assayed on poly[d(AT)] template.[23] Values are given relative to wild type (defined as 1.0).

[c] Relative to wild type (defined as 100 in both assays); note, however, that the reaction rates for wild-type Klenow fragment in the two assays are not the same. The substrates and assay methods were those described in the text. A value of 0.001 represents the lower limit of the assay using double-stranded DNA. A value of 0.4 is the lower limit of the single-stranded DNA assay so that it was not possible to assay those proteins having very low activity.

[d] These assays were not necessarily under V_{max} conditions, as judged by comparing the rates at different substrate concentrations.

[e] This value was calculated from the rate of degradation of a 5′-labeled duplex DNA. As indicated in the text, this assay is less sensitive than our standard duplex DNA assay. For comparison, the D355A protein, assayed at the same time, gave a very similar exonuclease rate, implying that the lower limit of this assay is around 0.01 (relative to wild type as 100).

[f] No reaction detectable after a 300-min incubation.

[g] Glu-357 is placed in both groups since the carboxylate side chain makes contacts both as a metal ligand and as a substrate-binding residue.

using phosphate buffers in the purification since we have found that some component used in our earlier procedure[22] (presumably the phosphate buffer) serves as source of pyrophosphate, resulting in a low level of Klenow fragment-catalyzed pyrophosphorysis of duplex DNA assay substrates, which can interfere with the assay of mutant proteins having very low 3′ → 5′-exonuclease activity.[7]

Characterization of Mutant Derivatives of Klenow Fragment

Measurement of Specific Activity of Polymerase

Polymerase activity was measured by the standard poly[d(A-T)] assay.[23] Protein concentrations were determined by the Bradford colorimetric assay,[24] using the reagent supplied by Bio-Rad. Either homogeneous Klenow fragment or bovine serum albumin (BSA) (of accurately determined concentration) has been used as the standard, with identical results. A polymerase-specific activity close to that of wild-type Klenow fragment (typically ~10^4 units/mg in this assay) was taken as evidence that the mutant proteins were not grossly misfolded. Because of the variability of the assay, it is less important that the polymerase-specific activity of the mutant protein have a particular numerical value than that it be similar to that of a wild-type standard assayed at the same time with the same reagents.

Exonuclease Assay on Double-Stranded DNA

1. Preparation of Assay Substrate.[25] The assay substrate was a heterogeneous mixture of restriction fragments carrying a single ^{32}P label at the 3′-terminal phosphodiester bond. The labeled substrate was prepared from *E. coli* chromosomal DNA digested to completion with *Sau*3AI. Digested DNA (24 μg, approximately 290 pmol of ends) was 3′-end-labeled using 1 unit of Klenow fragment in a 50-μl reaction containing 10 nmol unlabeled dGTP and 30 pmol [α-^{32}P]dATP (3000 Ci/mmol) for 10 min at room temperature. Excess unlabeled dATP (1 nmol) was added and incubation was continued for a further 1 min to ensure that all the 3′ ends were extended to the same extent (leaving a 2-nucleotide 5′ extension). The reaction was terminated by addition of EDTA to 20 m*M*, and the Klenow fragment was inactivated by heating at 70° for 15 min. The labeled DNA was phenol

[22] C. M. Joyce and N. D. F. Grindley, *Proc. Natl. Acad. Sci. USA* **80,** 1830 (1983).
[23] P. Setlow, *Methods. Enzymol.* **29,** 3 (1974).
[24] M. M. Bradford, *Anal. Biochem.* **72,** 248 (1976).
[25] P. S. Freemont, D. L. Ollis, T. A. Steitz, and C. M. Joyce, *Proteins* **1,** 66 (1986).

extracted, passed through a 1-ml c
unincorporated nucleotides, and rec
2. Assay Method.[8,25] The standard
M DNA 3′ termini in 6 m*M* Tris–
2-mercaptoethanol, and 50 m*M* NaCl
was incubated at 37°. Samples were
solution containing 1 mg/ml BSA (a
EDTA. The DNA was precipitated b
(w/v) trichloroacetic acid (TCA). Af
pelleted by spinning for 2 min in a
transferred to a fresh tube. The radioa
was determined by Cerenkov coun
quenching in the supernatant fractio
was added to both supernatant and
(prepared by mixing 0.5 ml each of the
of the counts observed in the two tu
fraction of the substrate ^{32}P that was
giving the rate of the exonuclease rea
number of polymerase units in the a
polymerase unit) of each mutant pro
that of wild type, which was arbitrarily
of the proteins in this study we were al
V_{max} because the same exonuclease r
higher substrate concentration.

Exonuclease Assay on Single-Stranded

1. Preparation of Assay Substrate. T
homopolymer, which was synthesized e
nucleotidyltransferase. An octanucleo
nmol) was incubated with 4 μmol [α-3
100-μl reaction containing 100 m*M* Tr
1 m*M* dithiothreitol (DTT), 10 m*M* Mg
nucleotidyltransferase for 16 hr at 37°. U
by gel filtration on a 1-ml Biogel P4 colu
mix was taken before and after the P4 c
filter and washed as described below. Co
filters gave the yield of labeled DNA
could be calculated the molarity of the as
per mole. The average chain length o
30 nucleotides) was determined by frac
polyacrylamide-urea sequencing gel foll

cation since we have found that some cedure[22] (presumably the phosphate hate, resulting in a low level of Klenow s of duplex DNA assay substrates, of mutant proteins having very low

tives of Klenow Fragment

Polymerase

l by the standard poly[d(A-T)] assay.[23] nined by the Bradford colorimetric l by Bio-Rad. Either homogeneous albumin (BSA) (of accurately deter- as the standard, with identical results. to that of wild-type Klenow fragment was taken as evidence that the mutant Because of the variability of the assay, e-specific activity of the mutant protein an that it be similar to that of a wild- ime with the same reagents.

ded DNA

.[25] The assay substrate was a heteroge- ts carrying a single ^{32}P label at the 3'- labeled substrate was prepared from to completion with *Sau*3AI. Digested nol of ends) was 3'-end-labeled using reaction containing 10 nmol unlabeled 0 Ci/mmol) for 10 min at room temper- nmol) was added and incubation was ure that all the 3' ends were extended eotide 5' extension). The reaction was 20 m*M*, and the Klenow fragment was 5 min. The labeled DNA was phenol

Natl. Acad. Sci. USA **80,** 1830 (1983).
1976).
nd C. M. Joyce, *Proteins* **1,** 66 (1986).

extracted, passed through a 1-ml column of Sephadex G-50 to remove unincorporated nucleotides, and recovered by ethanol precipitation.

2. Assay Method.[8,25] The standard reaction (20 μl) contained ~3×10^{-7} *M* DNA 3' termini in 6 m*M* Tris–HCl, pH 7.5, 6 m*M* $MgCl_2$, 6 m*M* 2-mercaptoethanol, and 50 m*M* NaCl. Enzyme was added and the mixture was incubated at 37°. Samples were removed at intervals into 0.5 ml of a solution containing 1 mg/ml BSA (as a precipitation carrier) and 10 m*M* EDTA. The DNA was precipitated by addition of an equal volume of 10% (w/v) trichloroacetic acid (TCA). After 5 to 10 min on ice, the DNA was pelleted by spinning for 2 min in a microfuge and the supernatant was transferred to a fresh tube. The radioactivity in both supernatant and pellet was determined by Cerenkov counting. To correct for the additional quenching in the supernatant fraction, 10 μl of a ^{32}P-containing solution was added to both supernatant and pellet from a "blank" precipitation (prepared by mixing 0.5 ml each of the BSA and TCA solutions). The ratio of the counts observed in the two tubes gave the correction factor. The fraction of the substrate ^{32}P that was solubilized was plotted versus time, giving the rate of the exonuclease reaction, which was normalized to the number of polymerase units in the assay. The exonuclease activity (per polymerase unit) of each mutant protein was then expressed relative to that of wild type, which was arbitrarily set at 100 (see Table II). For most of the proteins in this study we were able to show that this value reflected V_{max} because the same exonuclease rate was observed with a threefold higher substrate concentration.

Exonuclease Assay on Single-Stranded DNA[8]

1. Preparation of Assay Substrate. The substrate was a ^{32}P-labeled DNA homopolymer, which was synthesized enzymatically using terminal deoxynucleotidyltransferase. An octanucleotide primer, $p(dA)_8$ (typically 80 nmol) was incubated with 4 μmol [α-^{32}P]dATP (5 to 10 μCi/μmol) in a 100-μl reaction containing 100 m*M* Tris–HCl, pH 7.5, 100 μg/ml BSA, 1 m*M* dithiothreitol (DTT), 10 m*M* $MgCl_2$ and 80 units of terminal deoxynucleotidyltransferase for 16 hr at 37°. Unincorporated dATP was removed by gel filtration on a 1-ml Biogel P4 column. A 2-μl sample of the reaction mix was taken before and after the P4 column and was applied to a DE81 filter and washed as described below. Comparison of the counts on the two filters gave the yield of labeled DNA from the P4 column, from which could be calculated the molarity of the assay substrate and the radioactivity per mole. The average chain length of the substrate (typically around 30 nucleotides) was determined by fractionation of a sample on a 10% polyacrylamide-urea sequencing gel followed by densitometric scanning

and integration of all peaks visible on the resulting autoradiograph. A similar substrate was made by extension of $p(dT)_8$ with dTTP.

2. Assay Method. The standard reaction (20 μl) contained ~1 × 10^{-4} *M* DNA 3′ termini in 50 m*M* Tris–HCl, pH 7.5, 8 m*M* $MgCl_2$. Reactions were initiated by addition of enzyme and were incubated at 37°. At intervals, 2-μl samples were removed and quenched in 53 μl of 30 m*M* EDTA. The radioactivity remaining in single-stranded DNA was determined by applying 50 μl of each quenched solution to a 2.5-cm-diameter DE81 filter (Whatman). The released [^{32}P]dAMP was removed by washing the filters three times by gentle agitation for 5 min in 0.3 *M* ammonium formate, pH 8.0, followed by two washes in 95% (v/v) ethanol and one wash in ether.[26] After air-drying, the radioactivity present on each filter was determined by scintillation counting in 5 ml of Optifluor (Packard), and was plotted as a function of time to give the exonuclease rate for each protein. As in the previous method, this was normalized to the number of polymerase units in the assay and expressed relative to wild type, which was arbitrarily set at 100. For most of the proteins assayed, we could show that the exonuclease rate corresponded to V_{max} since the same rate was observed with a threefold higher substrate concentration. This assay method has also been used in our investigation of the pH dependence of the 3′→5′-exonuclease reaction.[8]

3. Steady-State Kinetics. Measurement of k_{cat} and K_m for wild-type Klenow fragment was carried out using poly(dT), made as described above for poly(dA), except that the specific activity of the labeled nucleotide was approximately 100-fold higher because lower concentrations of the poly(dT) substrate were to be used. The reaction mix (100 μl) contained 7.4 × 10^{-9} *M* Klenow fragment and poly(dT) (1 × 10^{-7} to 2 × 10^{-6} *M* of 3′ ends) in 6 m*M* Tris–HCl, pH 7.5, 6 m*M* $MgCl_2$, 6 m*M* 2-mercaptoethanol, 50 m*M* NaCl, and 100 μg/ml BSA. Samples (5 μl) were removed at intervals during incubation at 37° and processed as described above. Initial rates, determined by a least-squares analysis, were used to generate a Lineweaver–Burk double-reciprocal plot, from which k_{cat} and K_m were calculated.

Assessment of Exonuclease Assay Methods

Of the two exonuclease assays described above, the method using the duplex substrate is the more sensitive for assaying mutant proteins having very low exonuclease activity (see footnotes to Table II). This greater sensitivity results from two features of the duplex DNA assay. One is the

[26] F. R. Bryant, K. A. Johnson, and S. J. Benkovic, *Biochemistry* **22**, 3537 (1983).

quantitation of both the dNMP pro
the single-stranded DNA assay, only
fore, at low extents of conversion, t
sponds to the difference between t
important difference is that only t
substrate is labeled, whereas the singl
tract of labeled residues. As a resul
will give a greater proportionate rel
substrate (though this greater sensit
linear time course; see below). Althou
modifications, the sensitivity of the
improved by having a shorter tract
separation method (such as thin-layer
tation of the released dNMP as wel
methods used by other workers follo
assays we have described, but may
substrates used or the methods for sep
and product. Examples can be found

The exonuclease assay methods th
the most appropriate, given the tech
this work was carried out. With subs
however, some methods that we had fo
more feasible. In particular, we had
separation and quantitation of produ
quantitation of the DNA bands by d
insufficiently accurate for a detailed kin
to quantitate gels of this type accurat
technology has meant that measureme
following the degradation of a 5′-labe
attractive alternative, and we are starti
in our studies of the exonuclease reac

Although we have not used the gel a
tion of mutant proteins, our experience
strengths and weaknesses compared w
analysis of the degradation of a 5′-lab
measuring low levels of 3′ → 5′-exonucl
II) since it is difficult to quantitate a sma
of a large excess of starting material.
consuming and therefore less well suite
numbers of mutant proteins. Advanta
well-defined substrates, the potential f
strates (single- or double-stranded D

on the resulting autoradiograph. A
ion of p(dT)$_8$ with dTTP.
eaction (20 μl) contained ~1 × 10^{-4}
Cl, pH 7.5, 8 m*M* $MgCl_2$. Reactions
nd were incubated at 37°. At intervals,
enched in 53 μl of 30 m*M* EDTA.
-stranded DNA was determined by
ion to a 2.5-cm-diameter DE81 filter
was removed by washing the filters
nin in 0.3 *M* ammonium formate, pH
v/v) ethanol and one wash in ether.[26]
esent on each filter was determined
Optifluor (Packard), and was plotted
nuclease rate for each protein. As in
alized to the number of polymerase
ve to wild type, which was arbitrarily
ssayed, we could show that the exo-
nce the same rate was observed with
tion. This assay method has also been
pH dependence of the 3′→5′-exo-

ement of k_{cat} and K_m for wild-type
ng poly(dT), made as described above
ic activity of the labeled nucleotide
because lower concentrations of the
The reaction mix (100 μl) contained
poly(dT) (1 × 10^{-7} to 2 × 10^{-6} *M* of
m*M* $MgCl_2$, 6 m*M* 2-mercaptoethanol,
mples (5 μl) were removed at intervals
sed as described above. Initial rates,
ysis, were used to generate a Line-
, from which k_{cat} and K_m were calcu-

ethods

escribed above, the method using the
e for assaying mutant proteins having
footnotes to Table II). This greater
of the duplex DNA assay. One is the

kovic, *Biochemistry* **22**, 3537 (1983).

quantitation of both the dNMP product and the remaining substrate. (In the single-stranded DNA assay, only the substrate is quantitated and, therefore, at low extents of conversion, the amount of product formed corresponds to the difference between two very large numbers.) The second important difference is that only the 3′-terminal residue of the duplex substrate is labeled, whereas the single-stranded substrate contains a sizable tract of labeled residues. As a result, a low level of exonuclease activity will give a greater proportionate release of radioactivity from the duplex substrate (though this greater sensitivity comes at the expense of a less linear time course; see below). Although we ourselves have not made these modifications, the sensitivity of the single-stranded DNA assay could be improved by having a shorter tract of labeled residues and by using a separation method (such as thin-layer chromatography) that allows quantitation of the released dNMP as well as the remaining substrate. Assay methods used by other workers follow the same general principles as the assays we have described, but may differ in the precise details of the substrates used or the methods for separation and quantitation of substrate and product. Examples can be found in the references cited in Table III.

The exonuclease assay methods that we have used were chosen as being the most appropriate, given the technology available to us at the time this work was carried out. With subsequent improvements in technology, however, some methods that we had found unsatisfactory have now become more feasible. In particular, we had rejected methods that required the separation and quantitation of products on sequencing gels because the quantitation of the DNA bands by densitometry of autoradiographs was insufficiently accurate for a detailed kinetic study. More recently, the ability to quantitate gels of this type accurately and easily using phosphorimage technology has meant that measurement of a 3′ → 5′-exonuclease rate by following the degradation of a 5′-labeled oligonucleotide has become an attractive alternative, and we are starting to use this approach increasingly in our studies of the exonuclease reaction.

Although we have not used the gel assay extensively for the characterization of mutant proteins, our experience to date has allowed us to assess its strengths and weaknesses compared with the assays described above. Gel analysis of the degradation of a 5′-labeled substrate is less sensitive for measuring low levels of 3′ → 5′-exonuclease activity (see footnotes to Table II) since it is difficult to quantitate a small amount of product in the presence of a large excess of starting material. The gel assay is also more time consuming and therefore less well suited for the routine screening of large numbers of mutant proteins. Advantages of the gel assay are the use of well-defined substrates, the potential for using a variety of different substrates (single- or double-stranded DNAs, oligonucleotides of different

TABLE III
MUTATIONS THAT HAVE BEEN STUDIED IN THE CONSERVED "EXO" SEQUENCE MOTIFS OF DNA POLYMERASES

Enzyme	Mutation	Effect of mutation[a]	
		Exonuclease activity	Other
Exo I (Asp-355) motif[b]			
ϕ29 DNA polymerase	D12A[c]	~10^3-fold decrease	Defective in strand displacement
	E14A[c]	~300-fold decrease	Defective in strand displacement
T4 DNA polymerase	D112A, E114A[d]	~10^4-fold decrease	Mutator *in vivo*
	D112N[e]	Not measured	Mutator *in vivo*
T5 DNA polymerase	D164A, E166A[f]	Not detectable	
T7 DNA polymerase	D5A, E7A[g]	~10^5-fold decrease	
E. coli DNA polymerase II	D155A[h]	10^4-fold decrease	
	E157A[h]	10^3-fold decrease	
	D155A, E157A[h]	2×10^3-fold decrease	
B. subtilis DNA polymerase III	E427A[i]	Not detectable	Large decrease in polymerase activity
	E427Q[i]	~10^3-fold decrease	10-fold decrease in polymerase activity
	G430E[i]	10-fold decrease	2-fold decrease in polymerase activity
T. litoralis ("Vent") DNA polymerase	D141A, E143A[j]	Not detectable	
Yeast DNA polymerase II	D290A, E292A[k]	≥100-fold decrease	Mutator, *in vivo*
Yeast DNA polymerase III	D321A[l]	Not measured	Mutator, *in vivo*
	D321V[l]	Not measured	Mutator, *in vivo*
	E323A[l]	Not measured	Mutator, *in vivo*
Yeast mitochondrial DNA polymerase (MIP1)	D171G[m]	≥100-fold decrease	Mutator, *in vivo;* polymerase less processive
Exo II (Asp-424) motif			
ϕ29 DNA polymerase	D66A[c]	~400-fold decrease	Defective in strand displacement
T4 DNA polymerase	D219A[n]	>10^7-fold decrease	Mutator *in vivo*
E. coli DNA polymerase II	D228A[h]	10^3-fold decrease	
Yeast DNA polymerase III	D405A[l]	Not detectable in crude extract	Mutator *in vivo*
Yeast mitochondrial DNA polymerase (MIP1)	D230A[m]	≥100-fold decrease	Mutator *in vivo;* polymerase less processive
Exo III (Asp-501) motif			
ϕ29 DNA polymerase	D169A[o]	10^3-fold decrease	Defective in strand displacement
	Y165C[p]	24-fold decrease	Defective in strand displacement
	Y165F[p]	13-fold decrease	Defective in strand displacement
T4 DNA polymerase	D324A[d]	~10^4-fold decrease	Mutator *in vivo*
	D324G[q]	100-fold decrease[r]	Mutator *in vivo*
E. coli DNA polymerase II	Y330F[h]	60-fold decrease	
	D334A[h]	50-fold decrease	
Yeast mitochondrial DNA polymerase (MIP1)	D347A[m]	~500-fold decrease	Mutator *in vivo;* polymerase less processive
	C344G[m]	~3-fold decrease	Weak mutator *in vivo*

[a] Relative to wild type. Except where noted, the polymerase activity was essentially the same as wild type.
[b] To facilitate comparison with the Klenow fragment results, the important carboxylate metal ligand present in each motif of the Klenow fragment sequence is noted.

	D155A, E157A[n]	2×10^5-fold decrease	
B. subtilis DNA polymerase III	E427A[i]	Not detectable	Large decrease in polymerase activity
	E427Q[i]	~10^3-fold decrease	10-fold decrease in polymerase activity
	G430E[i]	10-fold decrease	2-fold decrease in polymerase activity
T. litoralis ("Vent") DNA polymerase	D141A, E143A[j]	Not detectable	
Yeast DNA polymerase II	D290A, E292A[k]	≥100-fold decrease	Mutator, *in vivo*
Yeast DNA polymerase III	D321A[l]	Not measured	Mutator, *in vivo*
	D321V[l]	Not measured	Mutator, *in vivo*
	E323A[l]	Not measured	Mutator, *in vivo*
Yeast mitochondrial DNA polymerase (MIP1)	D171G[m]	≥100-fold decrease	Mutator, *in vivo;* polymerase less processive
Exo II (Asp-424) motif			
ϕ29 DNA polymerase	D66A[c]	~400-fold decrease	Defective in strand displacement
T4 DNA polymerase	D219A[n]	>10^7-fold decrease	Mutator *in vivo*
E. coli DNA polymerase II	D228A[h]	10^3-fold decrease	

Yeast DNA polymerase III	D405A[l]	Not detectable in crude extract	Mutator *in vivo*
Yeast mitochondrial DNA polymerase (MIP1)	D230A[m]	≥100-fold decrease	Mutator *in vivo;* polymerase less processive
Exo III (Asp-501) motif			
ϕ29 DNA polymerase	D169A[p]	10^3-fold decrease	Defective in strand displacement
	Y165C[p]	24-fold decrease	Defective in strand displacement
	Y165F[p]	13-fold decrease	Defective in strand displacement
T4 DNA polymerase	D324A[d]	~10^4-fold decrease	Mutator *in vivo*
	D324G[q]	100-fold decrease[r]	Mutator *in vivo*
E. coli DNA polymerase II	Y330F[h]	60-fold decrease	
	D334A[h]	50-fold decrease	
Yeast mitochondrial DNA polymerase (MIP1)	D347A[m]	~500-fold decrease	Mutator *in vivo;* polymerase less processive
	C344G[m]	~3-fold decrease	Weak mutator *in vivo*

[a] Relative to wild type. Except where noted, the polymerase activity was essentially the same as wild type.
[b] To facilitate comparison with the Klenow fragment results, the important carboxylate metal ligand present in each motif of the Klenow fragment sequence is noted.
[c] A. Bernad, L. Blanco, J. M. Lázaro, G. Martin, and M. Salas, *Cell* **59,** 219 (1989).
[d] L. J. Reha-Krantz and R. L. Nonay, *J. Biol. Chem.* **268,** 27,100 (1993).
[e] L. J. Reha-Krantz, *J. Mol. Biol.* **202,** 711 (1988).
[f] D. K. Chatterjee and A. J. Hughes, unpublished work (1991).
[g] S. S. Patel, I. Wong, and K. A. Johnson, *Biochemistry* **30,** 511 (1991).
[h] Y. Ishino, H. Iwasaki, I. Kato, and H. Shinagawa, *J. Biol. Chem.* **269,** 14,655 (1994).
[i] M. H. Barnes, R. A. Hammond, C. C. Kennedy, S. L. Mack, and N. C. Brown, *Gene* **111,** 43 (1992).
[j] H. Kong, R. B. Kucera, and W. E. Jack, *J. Biol. Chem.* **268,** 1965 (1993).
[k] A. Morrison, J. B. Bell, T. A. Kunkel, and A. Sugino, *Proc. Natl. Acad. Sci. USA* **88,** 9473 (1991).
[l] M. Simon, L. Giot, and G. Faye, *EMBO J.* **10,** 2165 (1991).
[m] F. Foury and S. Vanderstraeten, *EMBO J.* **11,** 2717 (1992).
[n] M. W. Frey, N. G. Nossal, T. L. Capson, and S. J. Benkovic, *Proc. Natl. Acad. Sci. USA* **90,** 2579 (1993).
[p] M. S. Soengas, J. A. Esteban, J. M. Lázaro, A. Bernad, M. A. Blasco, M. Salas, and L. Blanco, *EMBO J.* **11,** 4227 (1992).
[q] L. J. Reha-Krantz, S. Stocki, R. L. Nonay, E. Dimayuga, L. D. Goodrich, W. H. Konigsberg, and E. K. Spicer, *Proc. Natl. Acad. Sci. USA* **88,** 2417 (1991).
[r] Double mutation with E191A, which alone has little effect on exonuclease activity.

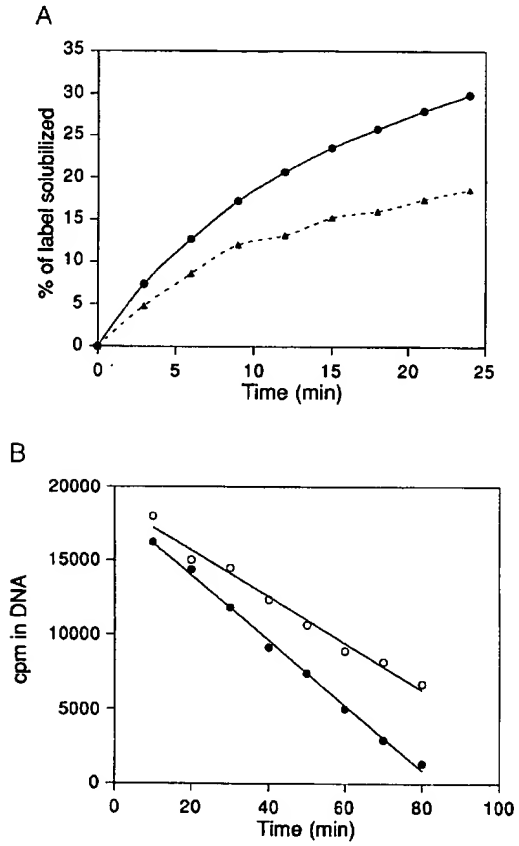


FIG. 2. Determination of 3′ →5 ′-exonuclease activity. (A) Degradation of a 3′-end-labeled duplex DNA by wild-type Klenow fragment (●) assayed at 1.3 n*M*, and the R455A mutant protein (▲) assayed at 3.6 n*M*. (B) Degradation of uniformly labeled single-stranded poly(dA), of average length 34, by wild-type Klenow fragment (●) assayed at 0.67 μM, and the L361A mutant protein (○) assayed at 0.87 μM. (C) Gel electrophoretic analysis of 3′ →5 ′-exonuclease activity, exemplified by the degradation of 5′-end-labeled $p(dT)_{14}$ by wild-type Klenow fragment. The reaction contained 6 μM oligonucleotide and 1.5 μM enzyme. Samples were removed at 15-sec intervals, as indicated. (D) Quantitation of the experiment shown in part (C). The extent of reaction (in μM) is presented either as the amount of 14-mer hydrolyzed (○), or as the number of phosphodiester bonds hydrolyzed (●) (calculated as described in the text). Early in the reaction, the two quantities are the same but, as the reaction proceeds, the second calculation method, which takes account of all species that can serve as substrates, is more satisfactory.

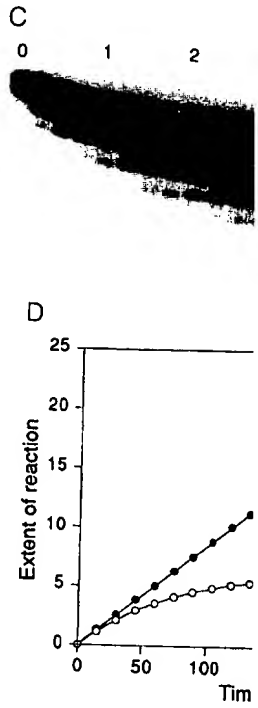
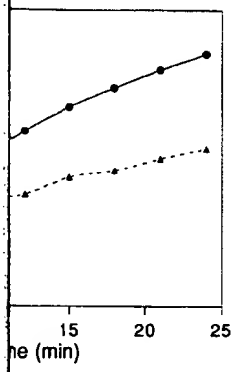
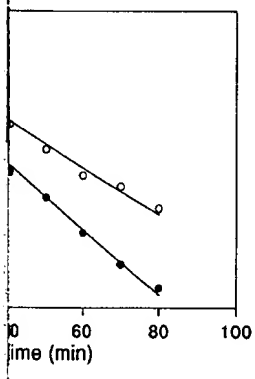


FIG. 2. (*co*

lengths, even extremely short oligonuc
the spectrum of reaction products as
the gel assay seems better suited for a
questions following the initial charact

When using a particular assay met
of how the reaction rate will change as
tial portion of the substrate is conver
which only the terminal residue is labe
strate) a single exonuclease event will
duce an unlabeled competing substr
only linear below 10 to 20% conversion
trast, the single-stranded substrate des
labeled residues so that the product fro
can serve as a substrate in subsequer
reaction rate remains constant for qui





e activity. (A) Degradation of a 3′-end-labeled
●) assayed at 1.3 n*M*, and the R455A mutant
of uniformly labeled single-stranded poly(dA),
ment (●) assayed at 0.67 μM, and the L361A
electrophoretic analysis of 3′ → 5′-exonuclease
nd-labeled $p(dT)_{14}$ by wild-type Klenow frag-
cleotide and 1.5 μM enzyme. Samples were
Quantitation of the experiment shown in part
d either as the amount of 14-mer hydrolyzed
s hydrolyzed (●) (calculated as described in
es are the same but, as the reaction proceeds,
ount of all species that can serve as substrates,

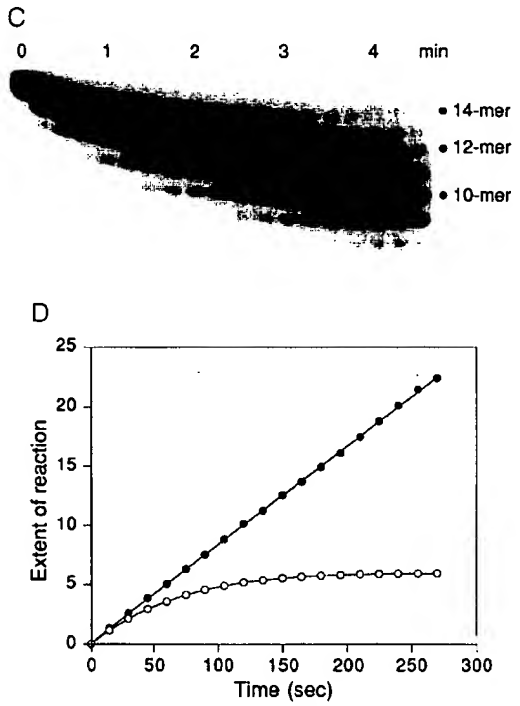


FIG. 2. (*continued*)

lengths, even extremely short oligonucleotides), and the ability to visualize the spectrum of reaction products as the reaction proceeds. As a result, the gel assay seems better suited for addressing more detailed mechanistic questions following the initial characterization of a mutant protein.

When using a particular assay method it is also important to be aware of how the reaction rate will change as the reaction proceeds and a substantial portion of the substrate is converted to product. With a substrate in which only the terminal residue is labeled (such as our duplex DNA substrate) a single exonuclease event will consume labeled substrate and produce an unlabeled competing substrate. Consequently, the reaction is only linear below 10 to 20% conversion of the substrate (Fig. 2A). By contrast, the single-stranded substrate described above contains more than 20 labeled residues so that the product from one round of exonuclease action can serve as a substrate in subsequent rounds, with the result that the reaction rate remains constant for quite a large extent of reaction (Fig.

2B). This same issue is nicely illustrated by the gel assay of the degradation of a 5′-labeled single-stranded DNA oligonucleotide (Fig. 2C). If one merely focuses on the rate of the first degradative event by measuring the rate of loss of the full-length substrate, then the time course will show substantial curvature as the reaction proceeds, due to the production of competing substrates. A more satisfactory approach is to consider the shorter species as potential substrates, as described by Cheng and Kuchta.[27] For this calculation, bands corresponding to substrate and all the reaction products are quantitated at each time point, and then the mole fraction of each species is multiplied by the number of exonuclease events required to generate that species, giving the amount of substrate degraded at each time point. Thus, for the degradation of 5′-^{32}P-labeled $(dT)_{14}$, the molar quantity of substrate degraded is given by {(fraction 14-mer)0 + (fraction 13-mer)1 + (fraction 12-mer)2 + (fraction 11-mer)3 + ...} × (moles of DNA in assay). Provided that the substrate is sufficiently long so that all the species under consideration are degraded at comparable rates, the reaction rate measured in this way remains linear for a substantial time (Fig. 2D).

Results and Interpretation of Mutational Studies of 3′ → 5′-Exonuclease of Klenow Fragment

Table II summarizes the assay results previously reported for Klenow fragment derivatives having mutations at the exonuclease active site,[7,8] together with some previously unreported data for mutations not included in our earlier study. As in any structure–function study involving data from mutant proteins, any meaningful interpretation of the data relies on the assumption that changes in protein structure due to the mutations are confined to the position of the altered side chain. Crystallographic studies of the single mutants D424A[7] and D355A[4] and of the D355A,E357A double mutant[7] validated the assumption for these proteins, and moreover suggested that the more conservative amide substitutions at these carboxylate positions would also have the same structure. For the other proteins in the study, we were able to draw on circumstantial evidence (wild-type levels of polymerase activity, similar overproduction yields and chromatographic behavior) that argued against gross structural perturbations in any of the mutant proteins. However, in the absence of further crystallographic data, we cannot rule out the possibility that some of the mutations may cause

[27] C.-H. Cheng and R. D. Kuchta, *Biochemistry* **32,** 8568 (1993).

subtle rearrangements within the acti
caveat in a study of this type.

Reaction Mechanism

The structural and mutational d
mechanism for the chemical step of
In this mechanism, the pair of di
apart, plays a pivotal role in the b
Metal A coordinates and polarizes t
nucleophilic displacement takes plac
oping negative charge on the pentaco
between the two metal sites. Metal

B

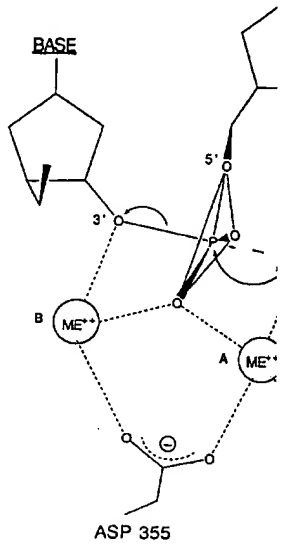


FIG. 3. The proposed transition state for the 3
is thought to involve catalysis mediated by the tv
ion A facilitates the formation of the attacking h
orientated toward the phosphorus by interaction
ion B stabilizes the geometry and charge of the
the departure of the 3′ hydroxyl group. Reproduc
Press, from L. Beese and T. A. Steitz, *EMBO J.*

d by the gel assay of the degradation oligonucleotide (Fig. 2C). If one degradative event by measuring the te, then the time course will show proceeds, due to the production of ctory approach is to consider the as described by Cheng and Kuchta.[27] ing to substrate and all the reaction point, and then the mole fraction of ber of exonuclease events required nount of substrate degraded at each of 5′-^{32}P-labeled $(dT)_{14}$, the molar n by {(fraction 14-mer)0 + (fraction ction 11-mer)3 + ...} × (moles of strate is sufficiently long so that all degraded at comparable rates, the mains linear for a substantial time

ional Studies of agment

ults previously reported for Klenow ns at the exonuclease active site,[7,8] rted data for mutations not included e–function study involving data from rpretation of the data relies on the structure due to the mutations are side chain. Crystallographic studies 5A[4] and of the D355A,E357A double these proteins, and moreover sug- de substitutions at these carboxylate ructure. For the other proteins in the mstantial evidence (wild-type levels oduction yields and chromatographic tructural perturbations in any of the nce of further crystallographic data, t some of the mutations may cause

32, 8568 (1993).

subtle rearrangements within the active-site region, and this is an important caveat in a study of this type.

Reaction Mechanism

The structural and mutational data together have led to a proposed mechanism for the chemical step of the exonuclease reaction[5,6] (Fig. 3). In this mechanism, the pair of divalent metal ions (A and B), 4 Å apart, plays a pivotal role in the bond making and breaking processes. Metal A coordinates and polarizes the attacking water molecule. As the nucleophilic displacement takes place, the metal ions stabilize the developing negative charge on the pentacovalent phosphorus center, positioned between the two metal sites. Metal B is also available to stabilize nega-

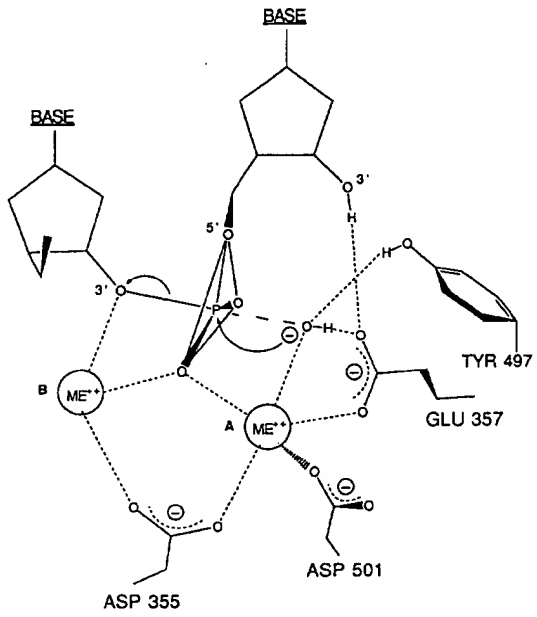


FIG. 3. The proposed transition state for the 3′ → 5′-exonuclease reaction. The mechanism is thought to involve catalysis mediated by the two bound divalent metal ions (ME^{++}). Metal ion A facilitates the formation of the attacking hydroxide ion, whose lone pair electrons are orientated toward the phosphorus by interactions with metal A, Tyr-497, and Glu-357. Metal ion B stabilizes the geometry and charge of the pentacovalent transition state and facilitates the departure of the 3′ hydroxyl group. Reproduced, with permission from Oxford University Press, from L. Beese and T. A. Steitz, *EMBO J.* **10**, 25 (1991).

tive charge on the leaving group and thus facilitate its departure. It has become increasingly apparent that two-metal-ion catalysis of this type may be a recurrent theme in phosphoryl transfer reactions,[28] so that the studies described here may have more generality than was at first supposed.

Although the mechanism of the chemical step of catalysis seems clear, we do not have a comparable understanding of the other steps that make up the kinetic pathway of the exonuclease reaction, nor do we know with certainty which step is rate limiting. It is clear from the steady-state kinetic parameters that there are substantial differences between the hydrolysis of single-stranded DNA and that of a duplex substrate. The exonucleolytic degradation of a duplex terminus is extremely slow ($k_{cat} \sim 10^{-3}$ sec^{-1}), but substrate binding appears to be very tight (K_m estimated to be in the nanomolar range).[29] For this and other reasons, it has been suggested that a duplex DNA substrate binds first to the polymerase site of Klenow fragment and is then transferred to the exonuclease site for hydrolysis, with the transfer step being rate limiting for the wild-type enzyme.[30] The degradation of single-stranded DNA is likely to be simpler kinetically since no transfer step is required, and the faster reaction ($k_{cat} \sim 0.1$ sec^{-1}) and higher K_m (5.6×10^{-7} M) are consistent with this expectation.[8] In either reaction, one cannot assume that the same step will be rate limiting when comparing a mutant enzyme with wild type, so that the measured change in exonuclease rate will not necessarily reflect the decrease in rate of a single step of the reaction. Bearing this caveat in mind, it is still possible to provide a structural rationale for the observed properties of the mutant proteins (Table II). It is probably reasonable to infer that mutations that weaken the binding of the active-site metal ions have caused the chemical step to become rate limiting; in these cases the extremely low exonuclease rates would reflect a failure of chemical catalysis. (Moreover, since the chemical step was not rate limiting for wild type, the decrease in the rate of this step as a result of the mutation must be in excess of the 10^4- to 10^5-fold decrease measured for the overall reaction rate.) Other mutations clearly have an effect because they remove an important substrate contact, even though one cannot say at present whether the effect of this loss is manifested kinetically in substrate binding (in the hydrolysis of single-stranded DNA), or transfer from the polymerase site (in the reaction with duplex DNA), or in the

[28] T. A. Steitz and J. A. Steitz, *Proc. Natl. Acad. Sci. USA* **90**, 6498 (1993).
[29] R. D. Kuchta, P. Benkovic, and S. J. Benkovic, *Biochemistry* **27**, 6716 (1988).
[30] C. E. Catalano, D. J. Allen, and S. J. Benkovic, *Biochemistry* **29**, 3612 (1990).

chemical step. Since some mutation
depending on whether the substrate
that the processes involved in bring
site are not entirely analogous in th
additional requirement for melting

Mutations in Carboxylate Ligands to

Figure 1 shows the details of the
at the exonuclease active site. Metal
coordination by one phosphate oxyge
355, Glu-357, and Asp-501, with a wa
nucleophile) as a fifth ligand. Metal I
phosphate oxygens, Asp-355 (shared
acting, via bridging water molecules, a
studies of three different mutant prot
ligand results in failure to bind one
mechanistic standpoint, two crystallog
cant. The D424A mutant protein bind
in an apparently normal manner, but
the D355A mutant protein binds subs
Because both the D355A and D424A
exonuclease activity (Table II), these
ment of both metal ions in catalysis.

The dramatic effect on exonucleas
site carboxylates is consistent with th
metal ions. Additional inferences can
detail. Mutations at Glu-357 are less
carboxylates, suggesting that Glu-357
The more important function of Glu-
made by this residue that contribute to
and the attacking nucleophile (see
substitutions at Asp-355, Asp-424, a
with the detailed coordination descri
355 and Asp-424 use both of the carb
tion, and therefore the Asn substitutio
By contrast, Asp-501 uses only one oxy
mutation has very little effect on the ex
E357A,D501N mutation causes a much
than would be expected from the con

nd thus facilitate its departure. It has
t two-metal-ion catalysis of this type
phoryl transfer reactions,[28] so that the
more generality than was at first sup-

chemical step of catalysis seems clear,
rstanding of the other steps that make
onuclease reaction, nor do we know
niting. It is clear from the steady-state
substantial differences between the
and that of a duplex substrate. The
uplex terminus is extremely slow (k_{cat}
appears to be very tight (K_m estimated
or this and other reasons, it has been
trate binds first to the polymerase site
ransferred to the exonuclease site for
being rate limiting for the wild-type
e-stranded DNA is likely to be simpler
required, and the faster reaction (k_{cat}
$\times 10^{-7}$ M) are consistent with this
ne cannot assume that the same step
ing a mutant enzyme with wild type,
exonuclease rate will not necessarily
ngle step of the reaction. Bearing this
o provide a structural rationale for the
nt proteins (Table II). It is probably
that weaken the binding of the active-
hemical step to become rate limiting;
konuclease rates would reflect a failure
since the chemical step was not rate
e in the rate of this step as a result of
the 10^4- to 10^5-fold decrease measured
ther mutations clearly have an effect
nt substrate contact, even though one
ffect of this loss is manifested kinetically
sis of single-stranded DNA), or transfer
reaction with duplex DNA), or in the

Acad. Sci. USA **90**, 6498 (1993).
nkovic, *Biochemistry* **27**, 6716 (1988).
nkovic, *Biochemistry* **29**, 3612 (1990).

chemical step. Since some mutations have quantitatively different effects depending on whether the substrate is single or double stranded, it seems that the processes involved in bringing the substrate to the exonuclease site are not entirely analogous in the two cases, perhaps because of the additional requirement for melting of a duplex terminus.

Mutations in Carboxylate Ligands to the Metal Ions

Figure 1 shows the details of the coordination of the two metal ions at the exonuclease active site. Metal A is bound in distorted tetrahedral coordination by one phosphate oxygen and the carboxylate groups of Asp-355, Glu-357, and Asp-501, with a water molecule (the proposed attacking nucleophile) as a fifth ligand. Metal B has octahedral coordination to two phosphate oxygens, Asp-355 (shared with metal A) and Asp-424, the latter acting, via bridging water molecules, as a bidendate ligand. Crystallographic studies of three different mutant proteins have shown that loss of a metal ligand results in failure to bind one or more of the metal ions.[4,7] From a mechanistic standpoint, two crystallographic results are particularly significant. The D424A mutant protein binds metal A and substrate (or product) in an apparently normal manner, but fails to bind metal B.[5-7] Conversely, the D355A mutant protein binds substrate and metal B but not metal A.[4] Because both the D355A and D424A mutant proteins have extremely low exonuclease activity (Table II), these data support the proposed involvement of both metal ions in catalysis.

The dramatic effect on exonuclease activity of mutations in the active-site carboxylates is consistent with the important role proposed for the metal ions. Additional inferences can be made by examining the data in detail. Mutations at Glu-357 are less severe than those at the other three carboxylates, suggesting that Glu-357 is the least important metal ligand. The more important function of Glu-357 may involve other interactions made by this residue that contribute to positioning the terminal nucleotide and the attacking nucleophile (see below). The results of asparagine substitutions at Asp-355, Asp-424, and Asp-501 are entirely consistent with the detailed coordination described for these residues. Both Asp-355 and Asp-424 use both of the carboxylate oxygens in metal coordination, and therefore the Asn substitution is not tolerated at these positions. By contrast, Asp-501 uses only one oxygen in metal binding and the D501N mutation has very little effect on the exonuclease activity. Interestingly, the E357A,D501N mutation causes a much greater loss in exonuclease activity than would be expected from the combination of the effects of the two

single mutations. This might be the consequence of losing two negative charges in the metal-binding region or, perhaps, Asp-501 can assist in binding the attacking nucleophile when Glu-357 is absent.

Mutations in Residues That Contact the Substrate

The effect on the exonuclease reaction of mutations in these residues is variable but quantitatively smaller than the effect of mutations in the metal ligands (Table II). We have therefore concluded that these residues play an important but less pivotal role than the metal ions and their ligands. The structural data suggest a probable role for these side chains in presenting the DNA substrate and the attacking nucleophile in the correct orientation for efficient catalysis. One of the most important residues in this category is likely to be Glu-357, whose carboxylate side chain is involved in a complex network of interactions.[6] One carboxylate oxygen interacts with metal A, while the other serves as a hydrogen-bond acceptor, both to the 3′ hydroxyl of the substrate and to the attacking water molecule. Given the involvement of both oxygens in these interactions, the severe effect of the glutamine substitution (E357Q) is as expected. Intriguingly, the E357A mutation seems better tolerated than E357Q, perhaps because the smaller alanine side chain allows access of a water molecule. The results of mutations at Tyr-497 confirm the importance of the observed hydrogen-bonding interaction between the phenolic hydroxyl and one of the oxygens of the terminal phosphodiester bond, since removal of the hydroxyl group alone (Y497F) has a similar effect to removal of the entire side chain (Y497A). The properties of the remaining mutations presumably reflect the varying degrees of importance of the interactions between the protein and the single-stranded terminal region of the DNA substrate. The stacking interaction between the terminal base and Phe-473 is clearly of primary importance. The Leu-361 residue is particularly interesting in that the L361A mutation has a much greater effect on the hydrolysis of duplex DNA than on hydrolysis of single-stranded DNA, implying that the intercalation of Leu-361 between the nucleotide bases at the 3′ terminus may be particularly important in the fraying that must accompany movement of a duplex substrate into the exonuclease site. The observed interactions between Gln-419 and Arg-455 and the phosphodiester backbone upstream of the point of hydrolysis (see Table I) appear to make very little contribution to the overall reaction, at least with the assay methods currently used. Clearly, however, the introduction of a negative charge in this region (Q419E) has severe consequences, presumably by interfering with DNA binding.

3′ → 5′-Exonuclease Active Site o

It is now clear that all DNA poly possess three small sequence motifs *et al.*[9] Conversely, in polymerases th are either completely absent (as in T catalytic residues (as in the eukaryot nal alignment of the three Exo mo work as it became apparent that a aligned incorrectly.[11,12] (This incorre sions in the studies of T4 DNA poly that had been mistakenly assigned to proficient proteins,[31] and led to a li exonuclease active site structures.[32]) T almost exactly the active site residue ture (Fig. 1; Table I), leading to the conservation reflects a similarity in t Exo I motif contains the core sequenc correspond to Asp-355 and Glu-357 sequence $NX_{2-3}(F/Y)D$; in Klenow Asp-424 and Asn-420, the latter intera of the 3′ terminus.[6] The Exo III moti the active site residues Tyr-497 and

Table III summarizes the results nuclease motifs of a number of DNA tions outside of the highly conserve consideration since, in the absence cannot be assessed.) The data of T. the proposal of a common active-si polymerases. Many of the exonuclease levels of polymerase activity, consiste exonuclease active sites that are struc Moreover, when detailed studies have constants for the polymerase reaction by mutations at the exonuclease site.[33]

[31] L. J. Reha-Krantz, S. Stocki, R. L. Nonay, E. I and E. K. Spicer, *Proc. Natl. Acad. Sci. USA*
[32] L. J. Reha-Krantz, *Gene* **112**, 133 (1992).
[33] B. T. Eger, R. D. Kuchta, S. S. Carroll, P. A. S. J. Benkovic, *Biochemistry* **30**, 1441 (1991).
[34] S. S. Patel, I. Wong, and K. A. Johnson, *Bioc*
[35] M. W. Frey, N. G. Nossal, T. L. Capson, and **90**, 2579 (1993).

the consequence of losing two negative
egion or, perhaps, Asp-501 can assist in
le when Glu-357 is absent.

tact the Substrate

e reaction of mutations in these residues
aller than the effect of mutations in the
e therefore concluded that these residues
role than the metal ions and their ligands.
obable role for these side chains in pre-
the attacking nucleophile in the correct
. One of the most important residues in
7, whose carboxylate side chain is involved
tions.[6] One carboxylate oxygen interacts
erves as a hydrogen-bond acceptor, both
ate and to the attacking water molecule.
oxygens in these interactions, the severe
ion (E357Q) is as expected. Intriguingly,
r tolerated than E357Q, perhaps because
ws access of a water molecule. The results
he importance of the observed hydrogen-
phenolic hydroxyl and one of the oxygens
ond, since removal of the hydroxyl group
fect to removal of the entire side chain
remaining mutations presumably reflect
e of the interactions between the protein
egion of the DNA substrate. The stacking
l base and Phe-473 is clearly of primary
ue is particularly interesting in that the
eater effect on the hydrolysis of duplex
-stranded DNA, implying that the interca-
ucleotide bases at the 3′ terminus may be
ing that must accompany movement of a
lease site. The observed interactions be-
the phosphodiester backbone upstream
ble I) appear to make very little contribu-
st with the assay methods currently used.
ion of a negative charge in this region
es, presumably by interfering with DNA

3′ → 5′-Exonuclease Active Site of Other DNA Polymerases

It is now clear that all DNA polymerases that have an editing function possess three small sequence motifs, named Exo I, II, and III by Bernad *et al.*[9] Conversely, in polymerases that do not have this function, the motifs are either completely absent (as in *Taq* DNA polymerase) or lack critical catalytic residues (as in the eukaryotic α DNA polymerases).[11,12] The original alignment of the three Exo motifs[9] has been modified in subsequent work as it became apparent that a few polymerase sequences had been aligned incorrectly.[11,12] (This incorrect alignment had significant repercussions in the studies of T4 DNA polymerase, since mutagenesis of residues that had been mistakenly assigned to the Exo I region yielded exonuclease-proficient proteins,[31] and led to a lively debate as to the universality of exonuclease active site structures.[32]) The three Exo sequence motifs parallel almost exactly the active site residues noted in the Klenow fragment structure (Fig. 1; Table I), leading to the obvious inference that the sequence conservation reflects a similarity in the active sites of these enzymes. The Exo I motif contains the core sequence DXE, where the two acidic residues correspond to Asp-355 and Glu-357 of Klenow fragment. Exo II has the sequence $NX_{2\text{-}3}(F/Y)D$; in Klenow fragment the conserved residues are Asp-424 and Asn-420, the latter interacting with the substrate just upstream of the 3′ terminus.[6] The Exo III motif has the sequence YX_3D, containing the active site residues Tyr-497 and Asp-501 in Klenow fragment.

Table III summarizes the results of mutations in the conserved exonuclease motifs of a number of DNA polymerases. (For simplicity, mutations outside of the highly conserved residues have been omitted from consideration since, in the absence of structural data, their significance cannot be assessed.) The data of Table III provide strong support for the proposal of a common active-site architecture for all proofreading polymerases. Many of the exonuclease-deficient derivatives have wild-type levels of polymerase activity, consistent with the idea of polymerase and exonuclease active sites that are structurally independent of one another. Moreover, when detailed studies have been carried out, individual kinetic constants for the polymerase reaction have been found to be unaffected by mutations at the exonuclease site.[33-35] In a few enzymes (herpes simplex

[31] L. J. Reha-Krantz, S. Stocki, R. L. Nonay, E. Dimayuga, L. D. Goodrich, W. H. Konigsberg, and E. K. Spicer, *Proc. Natl. Acad. Sci. USA* **88,** 2417 (1991).
[32] L. J. Reha-Krantz, *Gene* **112,** 133 (1992).
[33] B. T. Eger, R. D. Kuchta, S. S. Carroll, P. A. Benkovic, M. E. Dahlberg, C. M. Joyce, and S. J. Benkovic, *Biochemistry* **30,** 1441 (1991).
[34] S. S. Patel, I. Wong, and K. A. Johnson, *Biochemistry* **30,** 511 (1991).
[35] M. W. Frey, N. G. Nossal, T. L. Capson, and S. J. Benkovic, *Proc. Natl. Acad. Sci. USA* **90,** 2579 (1993).

virus DNA polymerase,[36] ϕ29 DNA polymerase,[37] *Bacillus subtilis* DNA polymerase III,[38] and yeast mitochondrial DNA polymerase[39]), mutations in the exonuclease region have been found to influence some aspects of the polymerase reaction. The reasons for this are unclear at present.

Uses of Exonuclease-Deficient DNA Polymerases

Many of the mutations listed in Table III were made solely to test the hypothesis of a conserved exonuclease active site. Others were constructed, using the sequence alignments as a guide, to facilitate particular experiments or to provide a research tool with wider applications. For the T4 and T7 DNA polymerases, both of which have very active 3′ → 5′-exonucleases, removal of the exonuclease was a necessary prerequisite for detailed kinetic studies of the polymerase function.[34,35] Exonuclease-deficient derivatives of Klenow fragment have proved invaluable in studying reactions involving mispaired bases at the polymerase site.[40] Inactivation of the exonuclease is also necessary for many biophysical experiments investigating the interaction between a polymerase and its DNA substrate in the presence of catalytically important metal-ion cofactors.

The particular attributes of an exonuclease-deficient polymerase that may make it useful as a research tool in biochemical manipulations are well illustrated by the properties of exonuclease-deficient derivatives of T7 DNA polymerase, which have found widespread use in DNA sequencing.[41,42] Because exonuclease-deficient T7 DNA polymerase cannot carry out the "idling" reaction (the turnover of dNTPs to dNMPs resulting from repeated incorporation and exonucleolytic excision by a stalled polymerase), it is better able to carry out strand-displacement synthesis or to synthesize through regions of secondary structure. Moreover, analogs such as dideoxynucleotides or α-thionucleotides are stably incorporated, instead of being rapidly removed by the exonuclease (as is the case with wild-type T7 DNA polymerase). The differences in the behavior of wild-type and exonuclease-deficient enzymes is much more pronounced for polymerases

[36] J. S. Gibbs, K. Weisshart, P. Digard, A. de Bruynkops, and D. M. Coen, *Mol. Cell. Biol.* **11,** 4786 (1991).
[37] M. S. Soengas, J. A. Esteban, J. M. Lázaro, A. Bernad, M. A. Blasco, M. Salas, and L. Blanco, *EMBO J.* **11,** 4227 (1992).
[38] M. H. Barnes, R. A. Hammond, C. C. Kennedy, S. L. Mack, and N. C. Brown, *Gene* **111,** 43 (1992).
[39] F. Foury and S. Vanderstraeten, *EMBO J.* **11,** 2717 (1992).
[40] C. M. Joyce, X. C. Sun, and N. D. F. Grindley, *J. Biol. Chem.* **267,** 24,485 (1992).
[41] S. Tabor and C. C. Richardson, *Proc. Natl. Acad. Sci. USA* **84,** 4767 (1987).
[42] S. Tabor and C. C. Richardson, *J. Biol. Chem.* **264,** 6447 (1989).

such as the T7 enzyme, which have Klenow fragment with a slower exon ment can itself carry out strand-displa dideoxy- or α-thionucleotide termin quencing gels as an exonuclease-def Klenow fragment, however, an exonu erable in circumstances where it is de of nucleotides, for example, when tryi tide analog or isotopically labeled nu quantities. An important caveat whe zyme in a "filling-in" reaction is that may result in accumulation of prod than expected.[43]

Other applications of exonuclease- vantage of the inability of these enz primer terminus. Although not widel oped for mutagenesis by forced misin using an exonuclease-deficient polym well-suited for random mutagenesis o of particular genetic traits by allele-s inability of the exonuclease-deficient matched terminus or, under the chose mismatch.[45,46] Our understanding of t 3′ → 5′-exonuclease active site of DNA chapter) is key to the development of e biotechnology tools. Because of this kn to design mutations to make any pol therefore there are no restrictions on most appropriate characteristics for th

Acknowledgments

During this work we have benefitted imme exonuclease structure provided by Tom Steitz an Chen Sun for excellent technical assistance and t manuscript. This work was supported by the Na to Nigel D. F. Grindley).

[43] J. M. Clark, C. M. Joyce, and G. P. Beardsley,
[44] X. Liao and J. A. Wise, *Gene* **88,** 107 (1990).
[45] C. R. Newton, A. Graham, I. E. Heptinstall, S *Nucleic Acids Res.* **17,** 2503 (1989).
[46] G. Sarkar, J. Cassady, C. D. K. Bottema, and S.

NA polymerase,[37] *Bacillus subtilis* DNA hondrial DNA polymerase[39]), mutations been found to influence some aspects of sons for this are unclear at present.

DNA Polymerases

in Table III were made solely to test the ease active site. Others were constructed, guide, to facilitate particular experiments h wider applications. For the T4 and T7 have very active 3′ → 5′-exonucleases, necessary prerequisite for detailed kinetic on.[34,35] Exonuclease-deficient derivatives invaluable in studying reactions involving e site.[40] Inactivation of the exonuclease ical experiments investigating the interac- DNA substrate in the presence of catalyti- rs.

n exonuclease-deficient polymerase that tool in biochemical manipulations are f exonuclease-deficient derivatives of T7 ound widespread use in DNA sequenc- cient T7 DNA polymerase cannot carry over of dNTPs to dNMPs resulting from ucleolytic excision by a stalled polymer- rand-displacement synthesis or to synthe- y structure. Moreover, analogs such as eotides are stably incorporated, instead xonuclease (as is the case with wild-type rences in the behavior of wild-type and much more pronounced for polymerases

. de Bruynkops, and D. M. Coen, *Mol. Cell. Biol.*

Lázaro, A. Bernad, M. A. Blasco, M. Salas, and

Kennedy, S. L. Mack, and N. C. Brown, *Gene* **111,**

J. **11,** 2717 (1992).
rindley, *J. Biol. Chem.* **267,** 24,485 (1992).
Natl. Acad. Sci. USA **84,** 4767 (1987).
Chem. **264,** 6447 (1989).

such as the T7 enzyme, which have a very active exonuclease, than for Klenow fragment with a slower exonuclease. Thus, wild-type Klenow fragment can itself carry out strand-displacement synthesis, is unable to degrade dideoxy- or α-thionucleotide termini, and gives the same pattern on sequencing gels as an exonuclease-deficient derivative.[20] Even when using Klenow fragment, however, an exonuclease-deficient enzyme may be preferable in circumstances where it is desirable to eliminate wasteful turnover of nucleotides, for example, when trying to incorporate a particular nucleotide analog or isotopically labeled nucleotide that is available only in small quantities. An important caveat when using an exonuclease-deficient enzyme in a "filling-in" reaction is that removal of the exonuclease activity may result in accumulation of products that are one nucleotide longer than expected.[43]

Other applications of exonuclease-deficient DNA polymerases take advantage of the inability of these enzymes to excise a mismatched DNA primer terminus. Although not widely used, procedures have been developed for mutagenesis by forced misincorporation and mismatch extension using an exonuclease-deficient polymerase, a strategy that is particularly well-suited for random mutagenesis over a defined region.[44] The detection of particular genetic traits by allele-specific amplification is based on the inability of the exonuclease-deficient polymerase either to remove a mismatched terminus or, under the chosen reaction conditions, to extend the mismatch.[45,46] Our understanding of the structure and mechanism of the 3′ → 5′-exonuclease active site of DNA polymerases (as summarized in this chapter) is key to the development of exonuclease-deficient polymerases as biotechnology tools. Because of this knowledge, it should be a simple matter to design mutations to make any polymerase exonuclease deficient, and therefore there are no restrictions on choosing the polymerase with the most appropriate characteristics for the desired application.

Acknowledgments

During this work we have benefitted immensely from the insights into the 3′ → 5′-exonuclease structure provided by Tom Steitz and colleagues. We are also grateful to Xiaojun Chen Sun for excellent technical assistance and to Nigel Grindley for a critical reading of the manuscript. This work was supported by the National Institutes of Health (grant GM-28550 to Nigel D. F. Grindley).

[43] J. M. Clark, C. M. Joyce, and G. P. Beardsley, *J. Mol. Biol.* **198,** 123 (1987).
[44] X. Liao and J. A. Wise, *Gene* **88,** 107 (1990).
[45] C. R. Newton, A. Graham, I. E. Heptinstall, S. J. Powell, C. Summers, and N. Kalsheker, *Nucleic Acids Res.* **17,** 2503 (1989).
[46] G. Sarkar, J. Cassady, C. D. K. Bottema, and S. S. Sommer, *Anal. Biochem.* **186,** 64 (1990).

Please note that only upper-case letters are considered to be aligned.

Alignment (DIALIGN format):

```
Pfu              1   MILDVDYITE EGKPVIRLFK KENGKFKIEH DRTFRPYIYA LLRDDSKIEE
Tgo              1   MILDTDYITE DGKPVIRIFK KENGEFKIDY DRNFEPYIYA LLKDDSAIED
KOD              1   MILDTDYITE DGKPVIRIFK KENGEFKIEY DRTFEPYFYA LLKDDSAIEE
Vent             1   MILDTDYITK DGKPIIRIFK KENGEFKIEL DPHFQPYIYA LLKDDSAIEE
Deep             1   MILDADYITE DGKPIIRIFK KENGEFKVEY DRNFRPYIYA LLKDDSQIDE
JDF-3            1   MILDVDYITE NGKPVIRVFK KENGEFRIEY DREFEPYFYA LLRDDSAIEE

Pfu             51   VKKITGERHG KIVRIVDVEK VEKKFLGKPI TVWKLYLEHP QDVPTIREKV
Tgo             51   VKKITAERHG TTVRVVRAEK VKKKFLGRPI EVWKLYFTHP QDVPAIRDKI
KOD             51   VKKITAERHG TVVTVKRVEK VQKKFLGRPV EVWKLYFTHP QDVPAIRDKI
Vent            51   IKAIKGERHG KTVRVLDAVK VRKKFLGREV EVWKLIFEHP QDVPAMRGKI
Deep            51   VRKITAERHG KIVRIIDAEK VRKKFLGRPI EVWRLYFEHP QDVPAIRDKI
JDF-3           51   IKKITAERHG RVVKVKRAEK VKKKFLGRSV EVWVLYFTHP QDVPAIRDKI

                                                      DXE (exo I)
Pfu            101   REHPAVVDIF EYDIPFAKRY LIDKGLIPME GEEELKILAF DIETLYHEGE
Tgo            101   KEHPAVVDIY EYDIPFAKRY LIDKGLIPME GDEELKMLAF DIETLYHEGE
KOD            101   REHGAVIDIY EYDIPFAKRY LIDKGLVPME GDEELKMLAF DIQTLYHEGE
Vent           101   REHPAVVDIY EYDIPFAKRY LIDKGLIPME GDEELKLLAF DIETFYHEGD
Deep           101   REHSAVIDIF EYDIPFAKRY LIDKGLIPME GDEELKLLAF DIETLYHEGE
JDF-3          101   RKHPAVIDIY EYDIPFAKRY LIDKGLIPME GEEELKLMSF DIETLYHEGE

Pfu            151   EFGKGPIIMI SYADENEAKV ITWKNIDLPY VEVVSSEREM IKRFLRIIRE
Tgo            151   EFAEGPILMI SYADEEGARV ITWKNIDLPY VDVVSTEKEM IKRFLKVVKE
KOD            151   EFAEGPILMI SYADEEGARV ITWKNVDLPY VDVVSTEREM IKRFLRVVKE
Vent           151   EFGKGEIIMI SYADEEEARV ITWKNIDLPY VDVVSNEREM IKRFvQVVKE
Deep           151   EFAKGPIIMI SYADEEEAKV ITWKKIDLPY VEVVSSEREM IKRFLKVIRE
JDF-3          151   EFGTGPILMI SYADESEARV ITWKKIDLPY VEVVSTEKEM IKRFLRVVKE

                                  NX2-3FD (exo II)
Pfu            201   KDPDIIVTYN GDSFDFPYLA KRAEKLGIKL TIGRDGS--E PKMQRIGDMT
Tgo            201   KDPDVLITYN GDNFDFAYLK KRSEKLGVKF ILGREGS--E PKIQRMGDRF
KOD            201   KDPDVLITYN GDNFDFAYLK KRCEKLGINF ALGRDGS--E PKIQRMGDRF
Vent           201   KDPDVIITYN GDNFDLPYLI KRAEKLGVRL VLGRDkehpE PKIQRMGDSF
Deep           201   KDPDVIITYN GDSFDLPYLV KRAEKLGIKL PLGRDGS--E PKMQRLGDMT
JDF-3          201   KDPDVLITYN GDNFDFAYLK KRCEKLGVSF TLGRDGS--E PKIQRMGDRF

Pfu            249   AVEVKGRIHF DLYHVITRTI NLPTYTLEAV YEAIFGKPKE KVYADEIAKA
Tgo            249   AVEVKGRIHF DLYPVIRRTI NLPTYTLEAV YEAIFGQPKE KVYAEEIAQA
KOD            249   AVEVKGRIHF DLYPVIRRTI NLPTYTLEAV YEAVFGQPKE KVYAEEITPA
```

```
Vent    251  AVEIKGRIHF DLFPVVRRTI NLPTYTLEAV YEAVLGKTKS KLGAEEIAAI
Deep    249  AVEIKGRIHF DLYHVIRRTI NLPTYTLEAV YEAIFGKPKE KVYAHEIAEA
JDF-3   249  AVEVKGRVHF DLYPVIRRTI NLPTYTLEAV YEAVFGKPKE KVYAEEIATA

                                  YX3D (exo III)
Pfu     299  WESGENLERV AKYSMEDAKA TYELGKEFLP MEIQLSRLVG QPLWDVSRSS
Tgo     299  WETGEGLERV ARYSMEDAKV TYELGKEFFP MEAQLSRLVG QSLWDVSRSS
KOD     299  WETGENLERV ARYSMEDAKV TYELGKEFLP MEAQLSRLIG QSLWDVSRSS
Vent    301  WETEESMKKL AQYSMEDARA TYELGKEFFP MEAELAKLIG QSVWDVSRSS
Deep    299  WETGKGLERV AKYSMEDAKV TYELGREFFP MEAQLSRLVG QPLWDVSRSS
JDF-3   299  WETGEGLERV ARYSMEDARV TYELGREFFP MEAQLSRLIG QGLWDVSRSS

Pfu     349  TGNLVEWFLL RKAYERNEVA PNKPSEEEYQ RRLRESYTGG FVKEPEKGLW
Tgo     349  TGNLVEWFLL RKAYERNELA PNKPDERELA RR-RESYAGG YVKEPERGLW
KOD     349  TGNLVEWFLL RKAYERNELA PNKPDEKELA RR-RQSYEGG YVKEPERGLW
Vent    351  TGNLVEWYLL RVAYARNELA PNKPDEEEYK RRLRTTYLGG YVKEPEKGLW
Deep    349  TGNLVEWYLL RKAYERNELA PNKPDEREYE RRLRESYAGG YVKEPEKGLW
JDF-3   349  TGNLVEWFLL RKAYERNELA PNKPDERELA RR-RggYAGG YVKEPERGLW

Pfu     399  ENIVYLDFRA LYPSIIITHN VSPDTLNLEG CKNYDIAPQV GHKFCKDIPG
Tgo     398  ENIVYLDFRS LYPSIIITHN VSPDTLNREG CEEYDVAPQV GHKFCKDFPG
KOD     398  ENIVYLDFRS LYPSIIITHN VSPDTLNREG CKEYDVAPQV GHRFCKDFPG
Vent    401  ENIIYLDFRS LYPSIIVTHN VSPDTLEKEG CKNYDVAPIV GYRFCKDFPG
Deep    399  EGLVSLDFRS LYPSIIITHN VSPDTLNREG CREYDVAPEV GHKFCKDFPG
JDF-3   398  DNIVYLDFRS LYPSIIITHN VSPDTLNREG CRSYDVAPEV GHKFCKDFPG

Pfu     449  FIPSLLGHLL EERQKIKTKM KETQDPIEKI LLDYRQKAIK LLANSFYGYY
Tgo     448  FIPSLLGDLL EERQKVKKKM KATIDPIEKK LLDYRQRAIK ILANSFYGYY
KOD     448  FIPSLLGDLL EERQKIKKKM KATIDPIERK LLDYRQRAIK ILANSYYGYY
Vent    451  FIPSILGDLI AMRQDIKKKM KSTIDPIEKK MLDYRQRAIK LLANSYYGYM
Deep    449  FIPSLLKRLL DERQEIKRKM KASKDPIEKK MLDYRQRAIK ILANSYYGYY
JDF-3   448  FIPSLLGNLL EERQKIKRKM KATLDPLEKN LLDYRQRAIK ILANSYYGYY

Pfu     499  GYAKARWYCK ECAESVTAWG RKYIELVWKE LEEKFGFKVL YIDTDGLYAT
Tgo     498  GYAKARWYCK ECAESVTAWG RQYIETTIRE IEEKFGFKVL YADTDGFFAT
KOD     498  GYARARWYCK ECAESVTAWG REYITMTIKE IEEKYGFKVI YSDTDGFFAT
Vent    501  GYPKARWYSK ECAESVTAWG RHYIEMTIRE IEEKFGFKVL YADTDGFYAT
Deep    499  GYAKARWYCK ECAESVTAWG REYIEFVRKE LEEKFGFKVL YIDTDGLYAT
JDF-3   498  GYARARWYCR ECAESVTAWG REYIEMVIRE LEEKFGFKVL YADTDGLHAT

Pfu     549  IPGGESEEIK KKALEFVKYI NSKLPGLLEL EYEGFYKRGF FVTKKRYAVI
Tgo     548  IPGADAETVK KKAKEFLDYI NAKLPGLLEL EYEGFYKRGF FVTKKKYAVI
```

```
KOD          548   IPGADAETVK KKAMEFLNYI NAKLPGALEL EYEGFYKRGF FVTKKKYAVI
Vent         551   IPGEKPELIK KKAKEFLNYI NSKLPGLLEL EYEGFYLRGF FVTKKRYAVI
Deep         549   IPGAKPEEIK KKALEFVDYI NAKLPGLLEL EYEGFYVRGF FVTKKKYALI
JDF-3        548   IPGADAETVK KKAMEFLNYI NPKLPGLLEL EYEGFYVRGF FVTKKKYAVI


Pfu          599   DEEGKVITRG LEIVRRDWSE IAKETQARVL ETILKHGDVE EAVRIVKEVI
Tgo          598   DEEDKITTRG LEIVRRDWSE IAKETQARVL EAILKHGDVE EAVRIVKEVT
KOD          598   DEEGKITTRG LEIVRRDWSE IAKETQARVL EALLKDGDVE KAVRIVKEVT
Vent         601   DEEGRITTRG LEVVRRDWSE IAKETQAKVL EAILKEGSVE KAVEVVRDVV
Deep         599   DEEGKIITRG LEIVRRDWSE IAKETQAKVL EAILKHGNVE EAVKIVKEVT
JDF-3        598   DEEGKITTRG LEIVRRDWSE IAKETQARVL EAILRHGDVE EAVRIVREVT


Pfu          649   QKLANYEIPP EKLAIYEQIT RPLHEYKAIG PHVAVAKKLA AKGVKIKPGM
Tgo          648   EKLSKYEVPP EKLVIYEQIT RDLKDYKATG PHVAVAKRLA ARGIKIRPGT
KOD          648   EKLSKYEVPP EKLVIHEQIT RDLKDYKATG PHVAVAKRLA ARGVKIRPGT
Vent         651   EKIAKYRVPL EKLVIHEQIT RDLKDYKAIG PHVAIAKRLA ARGIKVKPGT
Deep         649   EKLSKYEIPP EKLVIYEQIT RPLHEYKAIG PHVAVAKRLA ARGVKVRPGM
JDF-3        648   EKLSKYEVPP EKLVIHEQIT RELKDYKATG PHVAIAKRLA ARGVKIRPGT


Pfu          699   VIGYIVLRGD GPISNRAILA EEYDPKKHKY DAEYYIENQV LPAVLRILEG
Tgo          698   VISYIVLKGS GRIGDRAIPF DEFDPAKHKY DAEYYIENQV LPAVERILRA
KOD          698   VISYIVLKGS GRIGDRAIPF DEFDPTKHKY DAEYYIENQV LPAVERILRA
Vent         701   IISYIVLKGS GKISDRVILL TEYDPRKHKY DPDYYIENQV LPAVLRILEA
Deep         699   VIGYIVLRGD GPISKRAILA EEFDLRKHKY DAEYYIENQV LPAVLRILEA
JDF-3        698   VISYIVLKGS GRIGDRAIPF DEFDPTKHKY DADYYIENQV LPAVERILRA


Pfu          749   FGYRKEDLRY QKTRQVGLTS WLNIKKs--
Tgo          748   FGYRKEDLRY QKTRQVGLGA WLKPKt---
KOD          748   FGYRKEDLRY QKTRQVGLSA WLKPKGt--
Vent         751   FGYRKEDLRY QSSKQTGLDA WLKr-----
Deep         749   FGYRKEDLRW QKTKQTGLTA WLNIKKk--
JDF-3        748   FGYRKEDLRY QKTRQVGLGA WLKPKGkkk
```

Alignment (FASTA format):

```
>Pfu
MILDVDYITEEGKPVIRLFKKENGKFKIEHDRTFRPYIYALLRDDSKIEE
VKKITGERHGKIVRIVDVEKVEKKFLGKPITVWKLYLEHPQDVPTIREKV
REHPAVVDIFEYDIPFAKRYLIDKGLIPMEGEEELKILAFDIETLYHEGE
EFGKGPIIMISYADENEAKVITWKNIDLPYVEVVSSEREMIKRFLRIIRE
KDPDIIVTYNGDSFDFPYLAKRAEKLGIKLTIGRDGS--EPKMQRIGDMT
AVEVKGRIHFDLYHVITRTINLPTYTLEAVYEAIFGKPKEKVYADEIAKA
```

WESGENLERVAKYSMEDAKATYELGKEFLPMEIQLSRLVGQPLWDVSRSS
TGNLVEWFLLRKAYERNEVAPNKPSEEEYQRRLRESYTGGFVKEPEKGLW
ENIVYLDFRALYPSIIITHNVSPDTLNLEGCKNYDIAPQVGHKFCKDIPG
FIPSLLGHLLEERQKIKTKMKETQDPIEKILLDYRQKAIKLLANSFYGYY
GYAKARWYCKECAESVTAWGRKYIELVWKELEEKFGFKVLYIDTDGLYAT
IPGGESEEIKKKALEFVKYINSKLPGLLELEYEGFYKRGFFVTKKRYAVI
DEEGKVITRGLEIVRRDWSEIAKETQARVLETILKHGDVEEAVRIVKEVI
QKLANYEIPPEKLAIYEQITRPLHEYKAIGPHVAVAKKLAAKGVKIKPGM
VIGYIVLRGDGPISNRAILAEEYDPKKHKYDAEYYIENQVLPAVLRILEG
FGYRKEDLRYQKTRQVGLTSWLNIKKs--

>Tgo
MILDTDYITEDGKPVIRIFKKENGEFKIDYDRNFEPYIYALLKDDSAIED
VKKITAERHGTTVRVVRAEKVKKKFLGRPIEVWKLYFTHPQDVPAIRDKI
KEHPAVVDIYEYDIPFAKRYLIDKGLIPMEGDEELKMLAFDIETLYHEGE
EFAEGPILMISYADEEGARVITWKNIDLPYVDVVSTEKEMIKRFLKVVKE
KDPDVLITYNGDNFDFAYLKKRSEKLGVKFILGREGS--EPKIQRMGDRF
AVEVKGRIHFDLYPVIRRTINLPTYTLEAVYEAIFGQPKEKVYAEEIAQA
WETGEGLERVARYSMEDAKVTYELGKEFFPMEAQLSRLVGQSLWDVSRSS
TGNLVEWFLLRKAYERNELAPNKPDERELARR-RESYAGGYVKEPERGLW
ENIVYLDFRSLYPSIIITHNVSPDTLNREGCEEYDVAPQVGHKFCKDFPG
FIPSLLGDLLEERQKVKKKMKATIDPIEKKLLDYRQRAIKILANSFYGYY
GYAKARWYCKECAESVTAWGRQYIETTIREIEEKFGFKVLYADTDGFFAT
IPGADAETVKKKAKEFLDYINAKLPGLLELEYEGFYKRGFFVTKKKYAVI
DEEDKITTRGLEIVRRDWSEIAKETQARVLEAILKHGDVEEAVRIVKEVT
EKLSKYEVPPEKLVIYEQITRDLKDYKATGPHVAVAKRLAARGIKIRPGT
VISYIVLKGSGRIGDRAIPFDEFDPAKHKYDAEYYIENQVLPAVERILRA
FGYRKEDLRYQKTRQVGLGAWLKPKt---

>KOD
MILDTDYITEDGKPVIRIFKKENGEFKIEYDRTFEPYFYALLKDDSAIEE
VKKITAERHGTVVTVKRVEKVQKKFLGRPVEVWKLYFTHPQDVPAIRDKI
REHGAVIDIYEYDIPFAKRYLIDKGLVPMEGDEELKMLAFDIQTLYHEGE
EFAEGPILMISYADEEGARVITWKNVDLPYVDVVSTEREMIKRFLRVVKE
KDPDVLITYNGDNFDFAYLKKRCEKLGINFALGRDGS--EPKIQRMGDRF
AVEVKGRIHFDLYPVIRRTINLPTYTLEAVYEAVFGQPKEKVYAEEITPA
WETGENLERVARYSMEDAKVTYELGKEFLPMEAQLSRLIGQSLWDVSRSS
TGNLVEWFLLRKAYERNELAPNKPDEKELARR-RQSYEGGYVKEPERGLW
ENIVYLDFRSLYPSIIITHNVSPDTLNREGCKEYDVAPQVGHRFCKDFPG
FIPSLLGDLLEERQKIKKKMKATIDPIERKLLDYRQRAIKILANSYYGYY
GYARARWYCKECAESVTAWGREYITMTIKEIEEKYGFKVIYSDTDGFFAT
IPGADAETVKKKAMEFLNYINAKLPGALELEYEGFYKRGFFVTKKKYAVI
DEEGKITTRGLEIVRRDWSEIAKETQARVLEALLKDGDVEKAVRIVKEVT
EKLSKYEVPPEKLVIHEQITRDLKDYKATGPHVAVAKRLAARGVKIRPGT
VISYIVLKGSGRIGDRAIPFDEFDPTKHKYDAEYYIENQVLPAVERILRA
FGYRKEDLRYQKTRQVGLSAWLKPKGt--

>Vent
MILDTDYITKDGKPIIRIFKKENGEFKIELDPHFQPYIYALLKDDSAIEE
IKAIKGERHGKTVRVLDAVKVRKKFLGREVEVWKLIFEHPQDVPAMRGKI
REHPAVVDIYEYDIPFAKRYLIDKGLIPMEGDEELKLLAFDIETFYHEGD
EFGKGEIIMISYADEEEARVITWKNIDLPYVDVVSNEREMIKRFvQVVKE
KDPDVIITYNGDNFDLPYLIKRAEKLGVRLVLGRDkehpEPKIQRMGDSF
AVEIKGRIHFDLFPVVRRTINLPTYTLEAVYEAVLGKTKSKLGAEEIAAI
WETEESMKKLAQYSMEDARATYELGKEFFPMEAELAKLIGQSVWDVSRSS
TGNLVEWYLLRVAYARNELAPNKPDEEEYKRRLRTTYLGGYVKEPEKGLW
ENIIYLDFRSLYPSIIVTHNVSPDTLEKEGCKNYDVAPIVGYRFCKDFPG
FIPSILGDLIAMRQDIKKKMKSTIDPIEKKMLDYRQRAIKLLANSYYGYM
GYPKARWYSKECAESVTAWGRHYIEMTIREIEEKFGFKVLYADTDGFYAT

IPGEKPELIKKKAKEFLNYINSKLPGLLELEYEGFYLRGFFVTKKRYAVI
DEEGRITTRGLEVVRRDWSEIAKETQAKVLEAILKEGSVEKAVEVVRDVV
EKIAKYRVPLEKLVIHEQITRDLKDYKAIGPHVAIAKRLAARGIKVKPGT
IISYIVLKGSGKISDRVILLTEYDPRKHKYDPDYYIENQVLPAVLRILEA
FGYRKEDLRYQSSKQTGLDAWLKr-----

>Deep
MILDADYITEDGKPIIRIFKKENGEFKVEYDRNFRPYIYALLKDDSQIDE
VRKITAERHGKIVRIIDAEKVRKKFLGRPIEVWRLYFEHPQDVPAIRDKI
REHSAVIDIFEYDIPFAKRYLIDKGLIPMEGDEELKLLAFDIETLYHEGE
EFAKGPIIMISYADEEEAKVITWKKIDLPYVEVVSSEREMIKRFLKVIRE
KDPDVIITYNGDSFDLPYLVKRAEKLGIKLPLGRDGS--EPKMQRLGDMT
AVEIKGRIHFDLYHVIRRTINLPTYTLEAVYEAIFGKPKEKVYAHEIAEA
WETGKGLERVAKYSMEDAKVTYELGREFFPMEAQLSRLVGQPLWDVSRSS
TGNLVEWYLLRKAYERNELAPNKPDEREYERRLRESYAGGYVKEPEKGLW
EGLVSLDFRSLYPSIIITHNVSPDTLNREGCREYDVAPEVGHKFCKDFPG
FIPSLLKRLLDERQEIKRKMKASKDPIEKKMLDYRQRAIKILANSYYGYY
GYAKARWYCKECAESVTAWGREYIEFVRKELEEKFGFKVLYIDTDGLYAT
IPGAKPEEIKKKALEFVDYINAKLPGLLELEYEGFYVRGFFVTKKKYALI
DEEGKIITRGLEIVRRDWSEIAKETQAKVLEAILKHGNVEEAVKIVKEVT
EKLSKYEIPPEKLVIYEQITRPLHEYKAIGPHVAVAKRLAARGVKVRPGM
VIGYIVLRGDGPISKRAILAEEFDLRKHKYDAEYYIENQVLPAVLRILEA
FGYRKEDLRWQKTKQTGLTAWLNIKKk--

>JDF-3
MILDVDYITENGKPVIRVFKKENGEFRIEYDREFEPYFYALLRDDSAIEE
IKKITAERHGRVVKVKRAEKVKKKFLGRSVEVWVLYFTHPQDVPAIRDKI
RKHPAVIDIYEYDIPFAKRYLIDKGLIPMEGEEELKLMSFDIETLYHEGE
EFGTGPILMISYADESEARVITWKKIDLPYVEVVSTEKEMIKRFLRVVKE
KDPDVLITYNGDNFDFAYLKKRCEKLGVSFTLGRDGS--EPKIQRMGDRF
AVEVKGRVHFDLYPVIRRTINLPTYTLEAVYEAVFGKPKEKVYAEEIATA
WETGEGLERVARYSMEDARVTYELGREFFPMEAQLSRLIGQGLWDVSRSS
TGNLVEWFLLRKAYERNELAPNKPDERELARR-RggYAGGYVKEPERGLW
DNIVYLDFRSLYPSIIITHNVSPDTLNREGCRSYDVAPEVGHKFCKDFPG
FIPSLLGNLLEERQKIKRKMKATLDPLEKNLLDYRQRAIKILANSYYGYY
GYARARWYCRECAESVTAWGREYIEMVIRELEEKFGFKVLYADTDGLHAT
IPGADAETVKKKAMEFLNYINPKLPGLLELEYEGFYVRGFFVTKKKYAVI
DEEGKITTRGLEIVRRDWSEIAKETQARVLEAILRHGDVEEAVRIVREVT
EKLSKYEVPPEKLVIHEQITRELKDYKATGPHVAIAKRLAARGVKIRPGT
VISYIVLKGSGRIGDRAIPFDEFDPTKHKYDADYYIENQVLPAVERILRA
FGYRKEDLRYQKTRQVGLGAWLKPKGkkk

Sequence tree:
==========

Tree constructed using UPGMA

(((Pfu :0.000998,
Deep :0.000998):0.000080,
((Tgo :0.000905,
KOD :0.000905):0.000032,
JDF-3 :0.000937):0.000141):0.000067,
Vent :0.001144);

Please note that only upper-case letters are considered to be aligned.

Alignment (DIALIGN format):

```
Pfu              1   MILDVDYITE EGKPVIRLFK KENGKFKIEH DRTFRPYIYA LLRDDSKIEE
Tgo              1   MILDTDYITE DGKPVIRIFK KENGEFKIDY DRNFEPYIYA LLKDDSAIED
KOD              1   MILDTDYITE DGKPVIRIFK KENGEFKIEY DRTFEPYFYA LLKDDSAIEE
Vent             1   MILDTDYITK DGKPIIRIFK KENGEFKIEL DPHFQPYIYA LLKDDSAIEE
Deep             1   MILDADYITE DGKPIIRIFK KENGEFKVEY DRNFRPYIYA LLKDDSQIDE
JDF-3            1   MILDVDYITE NGKPVIRVFK KENGEFRIEY DREFEPYFYA LLRDDSAIEE

Pfu             51   VKKITGERHG KIVRIVDVEK VEKKFLGKPI TVWKLYLEHP QDVPTIREKV
Tgo             51   VKKITAERHG TTVRVVRAEK VKKKFLGRPI EVWKLYFTHP QDVPAIRDKI
KOD             51   VKKITAERHG TVVTVKRVEK VQKKFLGRPV EVWKLYFTHP QDVPAIRDKI
Vent            51   IKAIKGERHG KTVRVLDAVK VRKKFLGREV EVWKLIFEHP QDVPAMRGKI
Deep            51   VRKITAERHG KIVRIIDAEK VRKKFLGRPI EVWRLYFEHP QDVPAIRDKI
JDF-3           51   IKKITAERHG RVVKVKRAEK VKKKFLGRSV EVWVLYFTHP QDVPAIRDKI

Pfu            101   REHPAVVDIF EYDIPFAKRY LIDKGLIPME GEEELKILAF DIETLYHEGE
Tgo            101   KEHPAVVDIY EYDIPFAKRY LIDKGLIPME GDEELKMLAF DIETLYHEGE
KOD            101   REHGAVIDIY EYDIPFAKRY LIDKGLVPME GDEELKMLAF DIQTLYHEGE
Vent           101   REHPAVVDIY EYDIPFAKRY LIDKGLIPME GDEELKLLAF DIETFYHEGD
Deep           101   REHSAVIDIF EYDIPFAKRY LIDKGLIPME GDEELKLLAF DIETLYHEGE
JDF-3          101   RKHPAVIDIY EYDIPFAKRY LIDKGLIPME GEEELKLMSF DIETLYHEGE

Pfu            151   EFGKGPIIMI SYADENEAKV ITWKNIDLPY VEVVSSEREM IKRFLRIIRE
Tgo            151   EFAEGPILMI SYADEEGARV ITWKNIDLPY VDVVSTEKEM IKRFLKVVKE
KOD            151   EFAEGPILMI SYADEEGARV ITWKNVDLPY VDVVSTEREM IKRFLRVVKE
Vent           151   EFGKGEIIMI SYADEEEARV ITWKNIDLPY VDVVSNEREM IKRFvQVVKE
Deep           151   EFAKGPIIMI SYADEEEAKV ITWKKIDLPY VEVVSSEREM IKRFLKVIRE
JDF-3          151   EFGTGPILMI SYADESEARV ITWKKIDLPY VEVVSTEKEM IKRFLRVVKE

Pfu            201   KDPDIIVTYN GDSFDFPYLA KRAEKLGIKL TIGRDGS--E PKMQRIGDMT
Tgo            201   KDPDVLITYN GDNFDFAYLK KRSEKLGVKF ILGREGS--E PKIQRMGDRF
KOD            201   KDPDVLITYN GDNFDFAYLK KRCEKLGINF ALGRDGS--E PKIQRMGDRF
Vent           201   KDPDVIITYN GDNFDLPYLI KRAEKLGVRL VLGRDkehpE PKIQRMGDSF
Deep           201   KDPDVIITYN GDSFDLPYLV KRAEKLGIKL PLGRDGS--E PKMQRLGDMT
JDF-3          201   KDPDVLITYN GDNFDFAYLK KRCEKLGVSF TLGRDGS--E PKIQRMGDRF

Pfu            249   AVEVKGRIHF DLYHVITRTI NLPTYTLEAV YEAIFGKPKE KVYADEIAKA
Tgo            249   AVEVKGRIHF DLYPVIRRTI NLPTYTLEAV YEAIFGQPKE KVYAEEIAQA
KOD            249   AVEVKGRIHF DLYPVIRRTI NLPTYTLEAV YEAVFGQPKE KVYAEEITPA
```

```
Vent        251   AVEIKGRIHF DLFPVVRRTI NLPTYTLEAV YEAVLGKTKS KLGAEEIAAI
Deep        249   AVEIKGRIHF DLYHVIRRTI NLPTYTLEAV YEAIFGKPKE KVYAHEIAEA
JDF-3       249   AVEVKGRVHF DLYPVIRRTI NLPTYTLEAV YEAVFGKPKE KVYAEEIATA

Pfu         299   WESGENLERV AKYSMEDAKA TYELGKEFLP MEIQLSRLVG QPLWDVSRSS
Tgo         299   WETGEGLERV ARYSMEDAKV TYELGKEFFP MEAQLSRLVG QSLWDVSRSS
KOD         299   WETGENLERV ARYSMEDAKV TYELGKEFLP MEAQLSRLIG QSLWDVSRSS
Vent        301   WETEESMKKL AQYSMEDARA TYELGKEFFP MEAELAKLIG QSVWDVSRSS
Deep        299   WETGKGLERV AKYSMEDAKV TYELGREFFP MEAQLSRLVG QPLWDVSRSS
JDF-3       299   WETGEGLERV ARYSMEDARV TYELGREFFP MEAQLSRLIG QGLWDVSRSS

Pfu         349   TGNLVEWFLL RKAYERNEVA PNKPSEEEYQ RRLRESYTGG FVKEPEKGLW
Tgo         349   TGNLVEWFLL RKAYERNELA PNKPDERELA RR-RESYAGG YVKEPERGLW
KOD         349   TGNLVEWFLL RKAYERNELA PNKPDEKELA RR-RQSYEGG YVKEPERGLW
Vent        351   TGNLVEWYLL RVAYARNELA PNKPDEEEYK RRLRTTYLGG YVKEPEKGLW
Deep        349   TGNLVEWYLL RKAYERNELA PNKPDEREYE RRLRESYAGG YVKEPEKGLW
JDF-3       349   TGNLVEWFLL RKAYERNELA PNKPDERELA RR-RggYAGG YVKEPERGLW

                                DXXSLYPSII (Region II)
Pfu         399   ENIVYLDFRA LYPSIIITHN VSPDTLNLEG CKNYDIAPQV GHKFCKDIPG
Tgo         398   ENIVYLDFRS LYPSIIITHN VSPDTLNREG CEEYDVAPQV GHKFCKDFPG
KOD         398   ENIVYLDFRS LYPSIIITHN VSPDTLNREG CKEYDVAPQV GHRFCKDFPG
Vent        401   ENIIYLDFRS LYPSIIVTHN VSPDTLEKEG CKNYDVAPIV GYRFCKDFPG
Deep        399   EGLVSLDFRS LYPSIIITHN VSPDTLNREG CREYDVAPEV GHKFCKDFPG
JDF-3       398   DNIVYLDFRS LYPSIIITHN VSPDTLNREG CRSYDVAPEV GHKFCKDFPG

Pfu         449   FIPSLLGHLL EERQKIKTKM KETQDPIEKI LLDYRQKAIK LLANSFYGYY
Tgo         448   FIPSLLGDLL EERQKVKKKM KATIDPIEKK LLDYRQRAIK ILANSFYGYY
KOD         448   FIPSLLGDLL EERQKIKKKM KATIDPIERK LLDYRQRAIK ILANSYYGYY
Vent        451   FIPSILGDLI AMRQDIKKKM KSTIDPIEKK MLDYRQRAIK LLANSYYGYM
Deep        449   FIPSLLKRLL DERQEIKRKM KASKDPIEKK MLDYRQRAIK ILANSYYGYY
JDF-3       448   FIPSLLGNLL EERQKIKRKM KATLDPLEKN LLDYRQRAIK ILANSYYGYY

Pfu         499   GYAKARWYCK ECAESVTAWG RKYIELVWKE LEEKFGFKVL YIDTDGLYAT
Tgo         498   GYAKARWYCK ECAESVTAWG RQYIETTIRE IEEKFGFKVL YADTDGFFAT
KOD         498   GYARARWYCK ECAESVTAWG REYITMTIKE IEEKYGFKVI YSDTDGFFAT
Vent        501   GYPKARWYSK ECAESVTAWG RHYIEMTIRE IEEKFGFKVL YADTDGFYAT
Deep        499   GYAKARWYCK ECAESVTAWG REYIEFVRKE LEEKFGFKVL YIDTDGLYAT
JDF-3       498   GYARARWYCR ECAESVTAWG REYIEMVIRE LEEKFGFKVL YADTDGLHAT

Pfu         549   IPGGESEEIK KKALEFVKYI NSKLPGLLEL EYEGFYKRGF FVTKKRYAVI
Tgo         548   IPGADAETVK KKAKEFLDYI NAKLPGLLEL EYEGFYKRGF FVTKKKYAVI
```

```
KOD             548   IPGADAETVK KKAMEFLNYI NAKLPGALEL EYEGFYKRGF FVTKKKYAVI
Vent            551   IPGEKPELIK KKAKEFLNYI NSKLPGLLEL EYEGFYLRGF FVTKKRYAVI
Deep            549   IPGAKPEEIK KKALEFVDYI NAKLPGLLEL EYEGFYVRGF FVTKKKYALI
JDF-3           548   IPGADAETVK KKAMEFLNYI NPKLPGLLEL EYEGFYVRGF FVTKKKYAVI


Pfu             599   DEEGKVITRG LEIVRRDWSE IAKETQARVL ETILKHGDVE EAVRIVKEVI
Tgo             598   DEEDKITTRG LEIVRRDWSE IAKETQARVL EAILKHGDVE EAVRIVKEVT
KOD             598   DEEGKITTRG LEIVRRDWSE IAKETQARVL EALLKDGDVE KAVRIVKEVT
Vent            601   DEEGRITTRG LEVVRRDWSE IAKETQAKVL EAILKEGSVE KAVEVVRDVV
Deep            599   DEEGKIITRG LEIVRRDWSE IAKETQAKVL EAILKHGNVE EAVKIVKEVT
JDF-3           598   DEEGKITTRG LEIVRRDWSE IAKETQARVL EAILRHGDVE EAVRIVREVT


Pfu             649   QKLANYEIPP EKLAIYEQIT RPLHEYKAIG PHVAVAKKLA AKGVKIKPGM
Tgo             648   EKLSKYEVPP EKLVIYEQIT RDLKDYKATG PHVAVAKRLA ARGIKIRPGT
KOD             648   EKLSKYEVPP EKLVIHEQIT RDLKDYKATG PHVAVAKRLA ARGVKIRPGT
Vent            651   EKIAKYRVPL EKLVIHEQIT RDLKDYKAIG PHVAIAKRLA ARGIKVKPGT
Deep            649   EKLSKYEIPP EKLVIYEQIT RPLHEYKAIG PHVAVAKRLA ARGVKVRPGM
JDF-3           648   EKLSKYEVPP EKLVIHEQIT RELKDYKATG PHVAIAKRLA ARGVKIRPGT


Pfu             699   VIGYIVLRGD GPISNRAILA EEYDPKKHKY DAEYYIENQV LPAVLRILEG
Tgo             698   VISYIVLKGS GRIGDRAIPF DEFDPAKHKY DAEYYIENQV LPAVERILRA
KOD             698   VISYIVLKGS GRIGDRAIPF DEFDPTKHKY DAEYYIENQV LPAVERILRA
Vent            701   IISYIVLKGS GKISDRVILL TEYDPRKHKY DPDYYIENQV LPAVLRILEA
Deep            699   VIGYIVLRGD GPISKRAILA EEFDLRKHKY DAEYYIENQV LPAVLRILEA
JDF-3           698   VISYIVLKGS GRIGDRAIPF DEFDPTKHKY DADYYIENQV LPAVERILRA


Pfu             749   FGYRKEDLRY QKTRQVGLTS WLNIKKs--
Tgo             748   FGYRKEDLRY QKTRQVGLGA WLKPKt---
KOD             748   FGYRKEDLRY QKTRQVGLSA WLKPKGt--
Vent            751   FGYRKEDLRY QSSKQTGLDA WLKr-----
Deep            749   FGYRKEDLRW QKTKQTGLTA WLNIKKk--
JDF-3           748   FGYRKEDLRY QKTRQVGLGA WLKPKGkkk
```

Alignment (FASTA format):

```
>Pfu
MILDVDYITEEGKPVIRLFKKENGKFKIEHDRTFRPYIYALLRDDSKIEE
VKKITGERHGKIVRIVDVEKVEKKFLGKPITVWKLYLEHPQDVPTIREKV
REHPAVVDIFEYDIPFAKRYLIDKGLIPMEGEEELKILAFDIETLYHEGE
EFGKGPIIMISYADENEAKVITWKNIDLPYVEVVSSEREMIKRFLRIIRE
KDPDIIVTYNGDSFDFPYLAKRAEKLGIKLTIGRDGS--EPKMQRIGDMT
AVEVKGRIHFDLYHVITRTINLPTYTLEAVYEAIFGKPKEKVYADEIAKA
```

WESGENLERVAKYSMEDAKATYELGKEFLPMEIQLSRLVGQPLWDVSRSS
TGNLVEWFLLRKAYERNEVAPNKPSEEEYQRRLRESYTGGFVKEPEKGLW
ENIVYLDFRALYPSIIITHNVSPDTLNLEGCKNYDIAPQVGHKFCKDIPG
FIPSLLGHLLEERQKIKTKMKETQDPIEKILLDYRQKAIKLLANSFYGYY
GYAKARWYCKECAESVTAWGRKYIELVWKELEEKFGFKVLYIDTDGLYAT
IPGGESEEIKKKALEFVKYINSKLPGLLELEYEGFYKRGFFVTKKRYAVI
DEEGKVITRGLEIVRRDWSEIAKETQARVLETILKHGDVEEAVRIVKEVI
QKLANYEIPPEKLAIYEQITRPLHEYKAIGPHVAVAKKLAAKGVKIKPGM
VIGYIVLRGDGPISNRAILAEEYDPKKHKYDAEYYIENQVLPAVLRILEG
FGYRKEDLRYQKTRQVGLTSWLNIKKs--

>Tgo
MILDTDYITEDGKPVIRIFKKENGEFKIDYDRNFEPYIYALLKDDSAIED
VKKITAERHGTTVRVVRAEKVKKKFLGRPIEVWKLYFTHPQDVPAIRDKI
KEHPAVVDIYEYDIPFAKRYLIDKGLIPMEGDEELKMLAFDIETLYHEGE
EFAEGPILMISYADEEGARVITWKNIDLPYVDVVSTEKEMIKRFLKVVKE
KDPDVLITYNGDNFDFAYLKKRSEKLGVKFILGREGS--EPKIQRMGDRF
AVEVKGRIHFDLYPVIRRTINLPTYTLEAVYEAIFGQPKEKVYAEEIAQA
WETGEGLERVARYSMEDAKVTYELGKEFFPMEAQLSRLVGQSLWDVSRSS
TGNLVEWFLLRKAYERNELAPNKPDERELARR-RESYAGGYVKEPERGLW
ENIVYLDFRSLYPSIIITHNVSPDTLNREGCEEYDVAPQVGHKFCKDFPG
FIPSLLGDLLEERQKVKKKMKATIDPIEKKLLDYRQRAIKILANSFYGYY
GYAKARWYCKECAESVTAWGRQYIETTIREIEEKFGFKVLYADTDGFFAT
IPGADAETVKKKAKEFLDYINAKLPGLLELEYEGFYKRGFFVTKKKYAVI
DEEDKITTRGLEIVRRDWSEIAKETQARVLEAILKHGDVEEAVRIVKEVT
EKLSKYEVPPEKLVIYEQITRDLKDYKATGPHVAVAKRLAARGIKIRPGT
VISYIVLKGSGRIGDRAIPFDEFDPAKHKYDAEYYIENQVLPAVERILRA
FGYRKEDLRYQKTRQVGLGAWLKPKt---

>KOD
MILDTDYITEDGKPVIRIFKKENGEFKIEYDRTFEPYFYALLKDDSAIEE
VKKITAERHGTVVTVKRVEKVQKKFLGRPVEVWKLYFTHPQDVPAIRDKI
REHGAVIDIYEYDIPFAKRYLIDKGLVPMEGDEELKMLAFDIQTLYHEGE
EFAEGPILMISYADEEGARVITWKNVDLPYVDVVSTEREMIKRFLRVVKE
KDPDVLITYNGDNFDFAYLKKRCEKLGINFALGRDGS--EPKIQRMGDRF
AVEVKGRIHFDLYPVIRRTINLPTYTLEAVYEAVFGQPKEKVYAEEITPA
WETGENLERVARYSMEDAKVTYELGKEFLPMEAQLSRLIGQSLWDVSRSS
TGNLVEWFLLRKAYERNELAPNKPDEKELARR-RQSYEGGYVKEPERGLW
ENIVYLDFRSLYPSIIITHNVSPDTLNREGCKEYDVAPQVGHRFCKDFPG
FIPSLLGDLLEERQKIKKKMKATIDPIERKLLDYRQRAIKILANSYYGYY
GYARARWYCKECAESVTAWGREYITMTIKEIEEKYGFKVIYSDTDGFFAT
IPGADAETVKKKAMEFLNYINAKLPGALELEYEGFYKRGFFVTKKKYAVI
DEEGKITTRGLEIVRRDWSEIAKETQARVLEALLKDGDVEKAVRIVKEVT
EKLSKYEVPPEKLVIHEQITRDLKDYKATGPHVAVAKRLAARGVKIRPGT
VISYIVLKGSGRIGDRAIPFDEFDPTKHKYDAEYYIENQVLPAVERILRA
FGYRKEDLRYQKTRQVGLSAWLKPKGt--

>Vent
MILDTDYITKDGKPIIRIFKKENGEFKIELDPHFQPYIYALLKDDSAIEE
IKAIKGERHGKTVRVLDAVKVRKKFLGREVEVWKLIFEHPQDVPAMRGKI
REHPAVVDIYEYDIPFAKRYLIDKGLIPMEGDEELKLLAFDIETFYHEGD
EFGKGEIIMISYADEEEARVITWKNIDLPYVDVVSNEREMIKRFvQVVKE
KDPDVIITYNGDNFDLPYLIKRAEKLGVRLVLGRDkehpEPKIQRMGDSF
AVEIKGRIHFDLFPVVRRTINLPTYTLEAVYEAVLGKTKSKLGAEEIAAI
WETEESMKKLAQYSMEDARATYELGKEFFPMEAELAKLIGQSVWDVSRSS
TGNLVEWYLLRVAYARNELAPNKPDEEEYKRRLRTTYLGGYVKEPEKGLW
ENIIYLDFRSLYPSIIVTHNVSPDTLEKEGCKNYDVAPIVGYRFCKDFPG
FIPSILGDLIAMRQDIKKKMKSTIDPIEKKMLDYRQRAIKLLANSYYGYM
GYPKARWYSKECAESVTAWGRHYIEMTIREIEEKFGFKVLYADTDGFYAT

IPGEKPELIKKKAKEFLNYINSKLPGLLELEYEGFYLRGFFVTKKRYAVI
DEEGRITTRGLEVVRRDWSEIAKETQAKVLEAILKEGSVEKAVEVVRDVV
EKIAKYRVPLEKLVIHEQITRDLKDYKAIGPHVAIAKRLAARGIKVKPGT
IISYIVLKGSGKISDRVILLTEYDPRKHKYDPDYYIENQVLPAVLRILEA
FGYRKEDLRYQSSKQTGLDAWLKr-----

>Deep
MILDADYITEDGKPIIRIFKKENGEFKVEYDRNFRPYIYALLKDDSQIDE
VRKITAERHGKIVRIIDAEKVRKKFLGRPIEVWRLYFEHPQDVPAIRDKI
REHSAVIDIFEYDIPFAKRYLIDKGLIPMEGDEELKLLAFDIETLYHEGE
EFAKGPIIMISYADEEEAKVITWKKIDLPYVEVVSSEREMIKRFLKVIRE
KDPDVIITYNGDSFDLPYLVKRAEKLGIKLPLGRDGS--EPKMQRLGDMT
AVEIKGRIHFDLYHVIRRTINLPTYTLEAVYEAIFGKPKEKVYAHEIAEA
WETGKGLERVAKYSMEDAKVTYELGREFFPMEAQLSRLVGQPLWDVSRSS
TGNLVEWYLLRKAYERNELAPNKPDEREYERRLRESYAGGYVKEPEKGLW
EGLVSLDFRSLYPSIIITHNVSPDTLNREGCREYDVAPEVGHKFCKDFPG
FIPSLLKRLLDERQEIKRKMKASKDPIEKKMLDYRQRAIKILANSYYGYY
GYAKARWYCKECAESVTAWGREYIEFVRKELEEKFGFKVLYIDTDGLYAT
IPGAKPEEIKKKALEFVDYINAKLPGLLELEYEGFYVRGFFVTKKKYALI
DEEGKIITRGLEIVRRDWSEIAKETQAKVLEAILKHGNVEEAVKIVKEVT
EKLSKYEIPPEKLVIYEQITRPLHEYKAIGPHVAVAKRLAARGVKVRPGM
VIGYIVLRGDGPISKRAILAEEFDLRKHKYDAEYYIENQVLPAVLRILEA
FGYRKEDLRWQKTKQTGLTAWLNIKKk--

>JDF-3
MILDVDYITENGKPVIRVFKKENGEFRIEYDREFEPYFYALLRDDSAIEE
IKKITAERHGRVVKVKRAEKVKKKFLGRSVEVWVLYFTHPQDVPAIRDKI
RKHPAVIDIYEYDIPFAKRYLIDKGLIPMEGEEELKLMSFDIETLYHEGE
EFGTGPILMISYADESEARVITWKKIDLPYVEVVSTEKEMIKRFLRVVKE
KDPDVLITYNGDNFDFAYLKKRCEKLGVSFTLGRDGS--EPKIQRMGDRF
AVEVKGRVHFDLYPVIRRTINLPTYTLEAVYEAVFGKPKEKVYAEEIATA
WETGEGLERVARYSMEDARVTYELGREFFPMEAQLSRLIGQGLWDVSRSS
TGNLVEWFLLRKAYERNELAPNKPDERELARR-RggYAGGYVKEPERGLW
DNIVYLDFRSLYPSIIITHNVSPDTLNREGCRSYDVAPEVGHKFCKDFPG
FIPSLLGNLLEERQKIKRKMKATLDPLEKNLLDYRQRAIKILANSYYGYY
GYARARWYCRECAESVTAWGREYIEMVIRELEEKFGFKVLYADTDGLHAT
IPGADAETVKKKAMEFLNYINPKLPGLLELEYEGFYVRGFFVTKKKYAVI
DEEGKITTRGLEIVRRDWSEIAKETQARVLEAILRHGDVEEAVRIVREVT
EKLSKYEVPPEKLVIHEQITRELKDYKATGPHVAIAKRLAARGVKIRPGT
VISYIVLKGSGRIGDRAIPFDEFDPTKHKYDADYYIENQVLPAVERILRA
FGYRKEDLRYQKTRQVGLGAWLKPKGkkk

Sequence tree:
==========

Tree constructed using UPGMA

(((Pfu :0.000998,
Deep :0.000998):0.000080,
((Tgo :0.000905,
KOD :0.000905):0.000032,
JDF-3 :0.000937):0.000141):0.000067,
Vent :0.001144);